DOCUMENT RESUME

ED 149 293 CS 003 867

AUTHOR Carver, Ronald P.
TITLE Optimal Information Storage Rate for Reading Prose. Final Report.
INSTITUTION American Institutes for Research, Washington, D.C.; Missouri Univ., Kansas City.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
BUREAU NO 3-0267
PUB DATE 31 Dec 77
GRANT NEG-00-3-0043
NOTE 142p.; A number of pages may not reproduce well due to small type

EDRS PRICE MF-$0.83 HC-$7.35 Plus Postage.
DESCRIPTORS Cognitive Processes; Listening Skills; Post Secondary Education; Reading Ability; *Reading Comprehension; *Reading Processes; *Reading Rate; *Reading Research; *Reading Skills; *Retention; Speed Reading
IDENTIFIERS Rauding Theory

ABSTRACT

This is the final report on a research project that measured the effect of reading rate on accuracy and the efficiency of reading comprehension. At the outset of the project, existing research data were integrated into a comprehensive account called a "rauding theory," which interrelates reading ability, passage difficulty, reading time, and passage comprehension. With this theory as a foundation, this report consists of four separate studies: (1) Optimal Rate of Reading Prose, (2) Constancy of Reading Rate, (3) Reading Comprehension Under Normal and Speed Reading Conditions, and (4) Measuring Absolute Amounts of Reading Comprehension Using the Rauding Rescaling Procedure. Each study includes appropriate figures, tables, and a bibliography. (MAI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED149293

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Final Report

Project No. 3-0267
Grant No. NEG-00-3-0043

Optimal Information Storage Rate for Reading Prose

Ronald P. Carver
University of Missouri-Kansas City
Kansas City, Missouri 64110

December 31, 1977

The research reported herein was performed pursuant to a grant with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

U.S. DEPARTMENT OF

HEALTH, EDUCATION, AND WELFARE

NATIONAL INSTITUTE OF EDUCATION

CS003 867

FINAL REPORT

EXECUTIVE SUMMARY

(December 31, 1977)

Title: Optimal Information Storage Rate for Reading Prose

Project No. 3-0267

Grant No. NEG-00-3-0043

Principal Investigator: Dr. Ronald P. Carver

Institution: University of Missouri-Kansas City
5100 Rockhill Road
Kansas City, Missouri 64110

Project Duration: September 1, 1974 - December 31, 1977

Funding Schedule

American Institutes for Research*	University of Missouri-Kansas City
July 1, 1973 - Aug. 31, 1974	September 1, 1974 - December 31, 1977
$57,167	$52,325

*On August 31, 1974, the Principal Investigator moved from the American Institutes for Research in Washington, D.C. to the University of Missouri at Kansas City. A new grant was made to the University of Missouri to complete the project.

Project Description: The general purpose of the research was to investigate how reading accuracy, rate, and efficiency interrelate. Some theorists and some practioners contend that: (a) there is no such thing as an optimal rate of reading where efficiency is a maximum, and (b) good readers continually change rates. On the other hand there is some evidence to support a different viewpoint which holds that an individual's reading rate is relatively constant in typical or normal reading situations, and that forcing readers to proceed faster than this optimal rate will result in lower reading efficiency. Since those educators concerned with improving reading ability generally rely upon prevailing knowledge about reading, it seemed important that this knowledge be sound. Therefore, the present research was undertaken so as to provide more valid information regarding how rate affects the accuracy and efficiency of reading comprehension.

Project Results: At the outset of the project, existing research data were integrated into a comprehensive account of the interrelationships among reading time, reading rate, reading accuracy, reading efficiency, readability,

and reading ability. This extensive effort has been referred to as "rauding theory." Then, empirical data were collected relevant to this theory. In one study 102 college students were given passages to read at rates which varied from 83 to 500 words per minute, using motion picture film. At about 300 words per minute, the students were maximally efficient in their reading. This 300 words per minute was their optimal rate when they read material at grade 5, 8, 11, and 14 difficulty. Furthermore, when comparable passages were presented at the same listening rates, using time compressed speech techniques to vary listening rate, there was a similar optimal rate of about 300 words per minute. This means that these students had an optimal rate at which they comprehended passages, no matter what the level of difficulty of the material and no matter whether the passages were presented visually or auditorily. In another study of how fast college students normally read, it was found that this rate was around 300 words per minute. Thus, it appears that students normally read at a rate which is most efficient for them and that forcing them to read faster than this results in less efficiency of comprehension. This latter study involved a total of 666 students ranging from grade 3 through college. Each level of reading ability, read passages from grade 1 difficulty to grade 16 difficulty at a relatively constant rate. Another study varied the amount of time college students had to read passages so that on the average the rates varied from 62.5 to 100,000 words per minute. It was found that at the average rates of 1,000 words per minute or higher, comprehension scores were near chance or zero. The above studies also provided evidence to support some of the more precise relationships in rauding theory which are presented in the form of mathematical formulas. One of the inherent problems associated with this type of research involves the validity of the techniques used to measure comprehension. Also developed on this project was a technique for rescaling scores on comprehension tests so that they provide estimates of absolute amounts of reading comprehension. It is difficult, for example, to compare the results of one study which uses multiple choice questions as a measure of comprehension to another study which uses the number of passage blanks completed. The new technique developed on this project allows different types of comprehension measures to be directly compared.

Project Products. The theory of rauding has recently been published in the Reading Research Quarterly, Vol. 8, No. 1, 1977-78. There were two articles: "Toward a Theory of Reading Comprehension and Rauding" (pp. 8-63) and "Another Look at Rauding Theory," (pp. 116-132). Four other research articles prepared for publication are presented in this final report. These articles contain detailed reports of the research that was briefly outlined in the previous section. The titles of these reports are as follows: (a) "Optimal Rate of Reading Prose," (b) "Constancy of Reading Rate," (c) "Reading Comprehension Under Normal and Speed Reading Conditions," and (d) "Measuring Absolute Amounts of Reading Comprehension Using the Rauding Rescaling Procedure." Results of this project have been partially disseminated by presenting papers at the following annual meetings: (a) American Educational Research Association, 1974, (b) American Psychological Association, 1975, and (c) National Reading Conference, 1975 (subsequently published in 1975 NRC yearbook). At the 1976 annual meeting of the National Reading Conference an invited address was given which presented an overview of the results of this project.

Optimal Rate of Reading Prose

Ronald P. Carver

University of Missouri-Kansas City

September 1977

Optimal Rate of Reading Prose
Dr. Ronald P. Carver
10 W. Bridlespur Drive
Kansas City, Mo. 64114

Optimal Rate of Reading Prose[1]

Much of the basic research in prose reading has involved investigations of the cognitive processes involved (e.g., see Thorndyke, 1977). There is a sparsity of knowledge about the rate at which such information is processed. Current theoretical ideas about reading processes appear to be incompatible with empirical data regarding rate, accuracy, and efficiency of prose comprehension. Our understanding of the cognitive processes involved in reading should be enhanced by increases in basic empirical knowledge about how rate interacts with prose comprehension.

Relevant empirical data have been collected by Jester and Travers (1966). They investigated the effect of increased rate upon comprehension using both auditory and visual modes of prose presention. Based upon their data, and the auditory data of earlier researchers, they concluded that "the results clearly indicate an almost linear loss in comprehension as speed is increased [p. 301]." Jester and Travers also converted their scores into efficiency scores by dividing the number of items answered correctly by the time allowed, and they concluded, from these data that there was an optimal rate for the presentation of both auditory and visual prose material. For reading, they found an optimal rate of 300 words per minute and, for auding, they found an optimal rate of 200 words per min.

As for relevant cognitive theory, Neisser (1967) included a section on "Reading for Meaning" in his book Cognitive Psychology. In that section, he presents an argument, with little or no supporting empirical data, that individuals "read" at rates of 1200 words per min. or more. Implied in this presentation of Neisser's is the idea that there is no distinction to be made between reading and skipping, or skimming. Thus, the accuracy or efficiency accompanying the reading

process was largely disregarded, and there was no consideration of an optimal rate. Later theorists who have elaborated upon Neisser's ideas also seem to have discounted the idea that there might be a certain rate where reading efficiency is a maximum. These theorists are advocates of the hypothesis testing view of reading (Haber, 1974; Hochberg and Brooks, 1970; Smith, 1971). From this viewpoint, individuals are quite flexible in their rate because it purportedly varies with the nature of the information; the more redundant or predictable the letters or words or sentences, the faster the individual supposedly reads. Although the theory is not explicit, it would seem correct to infer that an individual changes rate so as to keep efficiency at a constantly high level across a wide range of rates. The esistence of an optimal rate for reading prose appears to be inconsistent with this viewpoint.

One of the primary purposes of the following research was to investigate more throughly the existence of an optimal rate during reading and auding. If such a phenomenon exists, then it would need to be accounted for in theory about the cognitive processes involved in reading.

A second primary purpose of the research was to collect empirical data relevant to a recently developed theory of reading comprehension, called rauding theory (Carver, 1977-78a). This theory is not explicit regarding the detailed cognitive processes involved in reading. However, it does hold that: (a) there is an optimal rate where efficiency is a maximum, (b) the optimal rate is the same during reading and auding, and (c) at the optimal rate, reading efficiency is equal to auding efficiency. As noted earlier, Jester and Travers found different optimal rates for reading and auding, 300 and 200 words per min., respectively. They also found different efficiency values at the optimal rates for reading and auding. Thus, further research is needed to determine whether the results of Jester and Travers provide reliable evidence against rauding theory.

There is another point of dissonance between rauding theory and the work Jester and Travers noted above. As mentioned earlier, the latter researchers contended that the relationship between rate and comprehension was linear. The earlier data collected by Goldstein (1940) and the more recent data presented by Carver (1973) also seems to suggest that as rate is increased, accuracy of comprehension decreases in an approximately linear manner. Yet, the relationship between rate and the accuracy of comprehension derived from rauding theory is hyperbolic, not linear, and there is one hyperbolic relationship for rates below the optimal rate and a completely different hyperbolic relationship at rates above the optimal rate. Thus, it is important to collect additional data relevant to the form of this functional relationship between rate and the accuracy of comprehension. The remainder of this introduction section will contain a recapitulation of rauding theory as it relates to optimal rate and the following research.

In rauding theory, the efficiency (E) of comprehending thoughts is defined as a product of accuracy (A) and rate (R),

$$E = AR. \quad (1)$$

Accuracy, A, is defined as the number of thoughts in a passage that were comprehended or understood per the number of thoughts presented in the passage. Conceptually, A has some properties that are similar to the more commonly used construct "precent comprehension." Rate, R, is defined as the number of thoughts presented per a unit of time. Conceptually, R has properties which are similar to the more familar "words per minute" measure. Efficiency, E, is therefore defined as the number of thoughts comprehended or understood per a unit of time. "Efficency" is a very familar concept in the area of reading, but E has a much more percise definition in rauding theory as compared to some of the more popular usages of the term, efficiency.

According to rauding theory, individuals require a certain amount of time to transform the words in a sentence into the complete thought that the author

intended to communicate. This theory is in keeping with the reading research of Poulton (1958) who, on the basis of his data, suggested that "...there is a limit to the amount of material which can be understood in a given time /p. 243/." Giving individuals more or less time than this threshold amount of time supposedly results in less than maximum efficiency of comprehension. The rate that is associated with the maximum efficiency is called the optimal rate. When the threshold amount of time is not allowed, i.e., as the rates of presentation increase beyond the optimal rate, then accuracy should drop preceptiously so that it is decreasing much more rapidly than rate is increasing; in turn, this means that the efficiency of understanding the thoughts decreases from the maximum. When the amount of time allowed for reading a passage is increased beyond the threshold amount of time, i.e., as the rate of presentation decreases from the optimal rate, then accuracy should increase because more time is allowed. However, accuracy should not increase in proportion to the decrease in rate. This is because "...the comprehension of a sentence is a one-trial learning event /p. 9, Carroll, 1968/," so that giving more time than is ordinarly necessary to comprehend each thought should result in diminishing returns with respect to the number of additional thoughts comprehended. This means that the efficiency of understanding the thoughts also should decrease from the maximum when rate decreases from the optimal rate.

As stated earlier, one of the primary purposes of the following research was to determine whether there is a rate of reading where efficiency is a maximum. In the presentation of rauding theory, Carver (1977-78a) presented direct evidence supporting Jester and Travers, as well as earlier researchers, in their contention that there is an optimal rate of auding; indirect evidence was also presented for an optimal rate of reading. The present research attempted to replicate the earlier auding results in a wider variety of situations, and it also provided direct evidence relevant to an optimal rate of reading. In the earlier research, instead

of directly varying the rate at which the words were presented, the amount of time allowed to read an entire passage was varied. Therefore, there was no direct evidence that efficiency peaked at the optimal rate because the words were not presented in an amount of time that was less than the threshold amount of time.

Another purpose of the following research was to investigate the constancy of reading rate. Rauding theory also holds that most individuals typically read at a constant rate, called their rauding rate, R_r. Individuals purportedly read at this constant rate because it is most efficient for them, i.e., the rauding rate is the optimal rate. Supposedly, individuals typically read at their capacity, i.e., at the highest rate at which they can transform the words in the sentence into the complete thoughts the author intended to communicate, and this rate is called the rauding rate. If they go faster or slower than the rauding rate they will be less efficient, so they read at the rauding rate to maximize efficiency. Since saying the words to oneself is considered to be the most efficient way of processing them, i.e., transforming the words in sentences into thoughts, (see Kleiman, 1975), rauding rate tends to be stabilized at a rate that is limited by a maximum internal speaking rate and also by a maximum cognitive processing rate.

If rauding theory is correct about rauding rate, R_r, being limited by a certain amount of time the individual needs to say the words in a sentence to himself or herself and to simultaneously formulate the thought out of the sentence, then there should be little difference between reading and auding for mature readers. This is because in the typical reading and auding situation the individual should be able to function at capacity. That capacity should be limited by central parameters associated with talking to oneself and processing the meaning of the words before it is limited by factors associated with inputting the visual or auditory stimulus. Thus, the highest efficiency associated with reading should be no higher than the highest efficiency associated with auding. Furthermore, the

optimal rate for reading should be no higher than the optimal rate for auding.

Another primary purpose of the research was to investigate the effect of different levels of the difficulty of the material upon the optimal rate. The data of Carver (1972), Miller and Coleman (1972), and Carver (1976) indicate that individuals tend to read material at a constant rate as long as rate is controlled for word length (e.g., by counting letters or syllables) and the difficulty level of the material does not exceed the ability level of the individual. Evidence from the following research that the optimal rate is the same for reading and auding, and is constant across difficulty levels, could be interpreted as providing further evidence to support the theoretical construct, rauding rate, R_r. Conversely, if the optimal rate varies with the mode of presentation and with the difficulty level of the material, then the validity of the rauding rate construct would be questionable.

A subsiderary research purpose involved the relationship between 1/E and 1/R. As defined earlier, efficiency, E, is the number of thoughts comprehended per a certain amount of time, and rate, R, is the number of thoughts presented per a certain amount of time. The inverse of efficiency, 1/E is the amount of time per each thought comprehended, and the inverse of rate, 1/R, is the amount of time per each thought presented. In rauding theory, the relationship between the amount of time per thought understood, 1/E, and the amount of time per thought presented, 1/R, is purported to be linear with a slope of 1.0 as long as the amount of time per sentence presented exceeds the threshold amount of time (Carver, 1977-78a). In equation form this relationship is:

$$\frac{1}{E} = (1.0)\,\frac{1}{R} + I. \qquad (2)$$

The I parameter in Equation 2 is a curve constant, called the "inefficiency constant," which represents the y intercept of the linear relationship between 1/E and 1/R. Notice that the slope of this straight line has been hypothesized to

be unity, 1.0. When this linear relationship between 1/E and 1/R in Equation 2 is algebraically transformed into the relationship between A and R (by substituting Equation 1 into Equation 2 and solving for A), the following function, which is recognized as hyperbolic, results:

$$A = \frac{1}{1.0 + IR} \qquad (3)$$

Equation 3 represents one of the two hyperbolic relationships between rate and comprehension that were alluded to earlier. Equation 3 holds for rates less than the optimal rate. Evidence relevant to the other hyperbolic relationship, for rates greater than the optimal rate, was not collected in the present study since that relationship is only valid when time is the primary variable being manipulated (see Carver, 1977-78b).

In reading theory, the relationship represented by Equation 2 is fundamental to the quantification of the theory. The following research investigated whether there was additional empirical support for the relationship represented by Equation 2. Evidence to support this Equation automatically provides evidence (a) for the hyperbolic relationship between A and R, and (b) against the linear relationship advocated by Jester and Travers (1966).

The most crucial factor involved in the conduct of this type of research is the estimation of the accuracy of comprehension or understanding. The research noted earlier (Carver, 1977-78a) used understanding judgements to estimate accuracy of comprehension. These judgments seemed to be more reliable and valid than the objective test scores that were used. The following research investigated further the validity of understanding judgements by comparing them to two different types of objective tests. One type of test, called the reading-storage test, RS-Test (Carver, 1974b; Carver, 1975a, Carver, 1975b), is completely objective in that it is developed objectively and scored objectively. The other type of test is a variant of the paraphrase test, P-Test (Carver, 1975b); the items were subjectively

developed but objectively scored. It was important to determine if the results of the research regarding optimal rate were dependent upon the type of measure used.

In summary, this research investigated: (a) whether individuals have an optimal rate of reading and an optimal rate of auding, (b) whether individuals have an optimal reading rate which is no different from their optimal auding rate, (c) whether the optimal rate is constant across levels of material difficulty, (d) whether the relationship between accuracy and rate for different levels of material difficulty can be represented by the general hyperbolic function of Equation 3, and (e) whether different methods of estimating the accuracy of comprehension or understanding tend to provide different results. Also, the results of this research will be interpreted with respect to certain specific predictions made from rauding theory regarding the accuracy and efficiency of comprehension at various rates.

Method

Introduction

Each subject participated in a full day of research. During the morning, there was a reading rate study, a preliminary reading and auding study that involved no tests, and two standardized tests. After lunch, there were two reading sessions and two auding sessions. Finally, there was a rapid reading study.

Three of the above data collection efforts were ancillary to the main research to be reported in this article. The first of these three was the initial reading rate study. It required the subjects to read six passages at their own rate for one minute each. There was no test involved, and the entire study required about 10 minutes. The second of the three involved the Raven Progressive Matrices Test. It was one of the two standardized tests mentioned above. The rapid reading study was the third data collection effort which was ancillary. The results of these three data collection efforts will be reported elsewhere.

Subjects

The 102 subjects were all college students who had responded to an advertisement in a university student newspaper. This advertisement stated that students would be paid $15.00 in cash, plus up to $13.00 in bonuses for their participation from 9:00 A.M. to 5:00 P.M. in an educational research project.

Procedure

The subjects were tested in 20 groups, ranging in size from 4 to 6 individuals. At the outset, the subjects were told the general nature of the research. They were informed that their bonus pay would depend upon how well they did on

the tests they took. After the short reading rate study, there was a series of six practice sessions for reading and listening. Each session involved 10 passages presented at different rates which each S was supposed to rate for understanding. After these six sessions, there was a 10 minute break, and then two forms of the National Reading Standards were administered. This is a newly developed reading test (Carver, 1977). Then, there was a 5 minute break which was followed by the administration of the Raven Progressive Matrices Test. After a 30 minute lunch break, there was a 15 minute practice period for the two reading and auding sessions that comprised the main research. Examples and practice in taking the two types of reading tests on the passages were given. During the two reading and two auding sessions, the subjects were presented a passage, after which they were asked to rate the percent of the passage that they understood. Then, they began working on a three minute test on the passage. There was an 8 minute break between each of the four sessions. After these four sessions, the rapid reading study was conducted.

The bonus to be paid for correct answers was explained to the subjects at the beginning of each session that involved tests. The reading rate study and the practice reading and auding sessions administered first involved no tests and, therefore, were the only parts of the entire research which involved no bonuses. The bonus system will be described in detail in a subsequent section.

Passages

There were a total of 80 different passages used. The passages were taken from the 330, 100-word passages that Bormuth (1969) sampled from curriculum materials in Grades 1-12 and college. As part of another study (Carver, 1975-76), the grade difficulty of the passages had been rated using the Rauding Scale of Prose Difficulty (Carver, 1974a). There were 20 passages chosen from each of four grade difficulties--5, 8, 11, and 14. The passages were selected randomly from the population of passages at each grade difficulty except any passage

that was being used in any other part of this research was excluded.

Design

There were 10 reading and listening rates--5, 10, 12.5, 15, 17.5, 20, 22.5, 25, 30, and ∞ standard length sentences per minute (Spm). By definition (Carver, 1977-78a) a standard length sentence consists of 16.7 standard length words, so that 15 Spm, for example, equals 250 standard length words per min. (Wpm). A standard length word contains 6 letter-spaces and a standard length sentence contains 100 letter-spaces, also by definition (Carver 1977-78a).

The 20 passages at each of the four grade difficulties were divided randomly into two sets, A and B, of 10 passages each. For each grade difficulty, 10 reading films and 10 auding tapes were made to control for passage differences and presentation order for the 10 rates used.

An example of the design for Grade 5 is presented in Table 1. The first column contains the 10 symbol codes for each of the 10 reading films, i.e., each row represents one film. Exactly the same code was used for the tapes, so each row also represents a tape. The succeeding 10 columns represent the order of presentation of the 10 passages. Within each cell of the table, the specific passage and specific rate of presentation is given. For example, the first film, or tape, is labeled 1-5A. The "1" indicates that it is the first film, or tape, in the series, the "5" indicates that it contains only material at fifth grade of difficulty, and the "A" indicates that only the Set A passages were used. For film (or tape) 1-5A, the symbols 5A9/20 are found under Column 4. The "5A9" means that the minth passage of the Set A passages at Grade 5 difficulty was presented. The "20" means that this passage was presented at a rate of 20 Spm. The fourth column means that the presentation order of this passage was fourth on this film (or tape). This same design for the reading films and the auding tapes was used for all four of the grade difficulties, i.e., not only for grade 5 but also for grades 8, 11, and 14.

Insert Table 1 about here

Optimal Rate

The above design provided control over possible order effects for the passages and the rates, as well as control over possible differences between passages within each grade difficulty.

Each of the 20 groups of subjects received two of the films and two of the tapes at two different grade difficulties. The first 10 groups received the Grade 8 and Grade 11 passages, and the second 10 groups received the Grade 5 and Grade 14 passages. Each group was presented one of the set of 10 passages (i.e., A or B) at each grade in the reading mode (a reading film), and the other set of 10 passages in the auding mode (an auding tape). For example, the first group tested received the 11A set of passages in the reading mode, the 11B set of passages in the auding mode, the 8A set of passages in the reading mode, and the 8B set of passages in the auding mode. The above design provided for control of individual differences when comparing reading and listening at each grade difficulty.

The first group tested received the two films before the two tapes. For the second group, this order was reversed. This switching was continued for all the 20 groups so that the odd numbered groups always received the two reading films first and the even numbered groups always received the two listening tapes first.

There were a total of 40 films and tapes at Grades 8 and 11 which were used during the 4 sessions for the first ten groups, and similarly a total of 40 films and tapes which were used for the second set of 10 groups. There was a partial counterbalancing of the films and tapes presented in each of the four sessions so that each difficulty level appeared equally often in each presentation order of the four sessions; complete counterbalancing was impossible under the preceding design constraints.

The above design controlled possible presentation order effects for reading versus auding and for one grade difficulty versus the other grade difficulty.

Out of each group of four to six subjects, the odd numbered subjects in each group were administered a reading-storage type of test and the even numbered subje were administered a paraphrase type of test. The nature of these two types of tes is explained in the next section.

Experimental Tests

A reading-storage test (RS-test) on each of the 80 passages was prepared usin the standard algorithm for the revised version of this type of test (Carver, 1975a This algorithm provided 10 test items for each passage, i.e., 1 item per each set consecutive 10 words. The reading-storage type of test is made from the passage i self by deleting the letters of every other word except the initial letter. Out o every five consecutive initial letters, one is replaced with a wrong letter. For test, an individual is asked to circle the wrong letters. The subjects were infor that they would receive 4¢ per each correct answer on each test item on each passa

A paraphrase type of test was also developed for each of the 80 passages. This type of test was somewhat similar to that used by Carver (1975b) here a multiple-choice item was developed for each sentence in a passage. However, in the earlier research, the item stem and the correct alternative were all developed using synonyms almost exclusively, i.e., almost none of the original structual words were used in the question. In the present study, the focus was entirely upon writing items that could not be answered correctly without the passage being presented, and could be answered correctly after the passage had been presented for a reasonable length of time, e.g., at a rate of about 150 words per minute. This latter focus allowed the exact words to be used in the correct answers as well as the alternate wrong answers. The advantage of the new focus in item development was that the exact words used in the passage could be used with the meaning changed to provide excellent distractors. Yet, if these types of distractors were used without ever being in the correct alternative then this would provide an unwanted clue indicating that the distractor could

not be correct since it contained some of the same words from the passage. Therefore, it was necessary that the restriction upon using synonyms be changed so that sometimes the correct alternative contained some of the exact same words or phrases used in the passage. There was one other difference between the present items and the earlier similar items and that was that the present items had four possible alternative answers to choose from whereas the earlier research used five distractors. The items were originally drafted by a research assistant, and later edited by the experimenter prior to their use.

The bonus system for these items was exactly the same as for the RS-test items.

Standardized Tests

The National Reading Standards is a five level standardized reading test. Each level requires 15 minutes to complete (Carver, 1977).. Each subject was administered one level, Test 5, of both forms of the test, i.e., Form A and Form B. However, one of the two forms was administered as a Short Form of the test, i.e., a 7.5 minute time limit was employed rather than the standard 15 minutes. The test provides a reading efficiency score (RE score) which can be transformed into a grade ability score (G_a), a level ability score (L_a), and a reading rate score (R_w). The grade ability score has been calibrated so that any particular grade ability score means that the individual is expected to be able to read and understand one-half of all the material at the grade difficulty, as measured by the Rauding Scale of Prose Difficulty, noted earlier. For example, if an individual receives a Grade 8 score on the NRS, this may be interpreted as meaning that the individual is expected to be able to read and understand approximately 50% of all the curriculum materials that are at Grade 8 difficulty according to the Rauding Scale. The level ability score consists of three grade abilities, i.e., Grades 1-3 ability are Level 1 ability, Grades 4-6 ability are Level 2 ability, etc. It may be noted that the four grade difficulties of

the experimental passages--5, 8, 11, and 14--were selected to represent the midpoint of Levels 2, 3, 4, and 5 difficulties.

The order of presentation of Form A and Form B, and the order of presentation of the 15 minute version and the 7.5 minute version was counterbalanced for each successive four groups of subjects. For example, Groups 1, 5, 9, 13, and 17 received Test 5, Form A, using the regular 15 minute time limit, and Test 5, Form B, using the shortened 7.5 minute time limit. The subjects were informed that they would received a bonus of 1¢ for each 5 correct answers on both the 15 and 7.5 minute versions, except a correction for guessing would be applied first on the 7.5 minute version.

Rate Variable

Earlier it was noted that there were 10 presentation rates--5, 10, 12.5, 15, 17.5, 20, 22.5, 25, 30 and ∞ Spm. The infinity rate, ∞, means that the test on the passage was administered without the passage being presented, i.e., the time allowed for reading was zero, thus producing an infinitly high rate of reading. As noted earlier, rate in Spm has been defined as 16.7 standard length words per minute, a standard length word contains 6 letter-spaces, and there are 100 letter-spaces per standard length sentence. Therefore, rate in sentences per min., Spm, is also the number of 100 character spaces per minute.

Rate in Spm was developed from the typed version of the passages. Each passage was typed on a piece of paper with a maximum margin of 60 spaces, 12 spaces per inch. No words were hyphenated at the margin, i.e., if a word was going to exceed the 60 space limit, then it was moved to the beginning of the following line. After all passages had been typed using these rules, the number of character spaces on the line was counted starting with the first letter in the first word on the line and ending with the blank space following the last word on the line. Dividing this number of character spaces by 100

provided the number of standard length sentences (S) on the line. The total number of such sentences, S, in the passage was determined by adding up the S values for all the lines.

For each auding passage, the total amount of time in minutes required to produce a 10 Spm rate was calculated by dividing the S value for the passage by 10 Spm. Then, a professional announcer practiced reading the passage until he was able to produce one that lasted this amount of time, i.e., the rate was therefore 10 Spm. The rates faster than 10 Spm, i.e., 12.5 to 30 Spm, were produced using a rate-controlled speech device (i.e., a Tempo-Regulator). This technique usually involved more than one trial in getting the passage compressed so that the actual time the passage was presented matched the time determined by dividing S by the particular Spm involved. The 5 Spm rate is an average rate that was produced by playing the 10 Spm passage twice. Each particular passage was presented at 5 of the total of 10 possible rates. After the prescribed rate for each passage was developed, they were spliced together in the prescribed order to produce the total of 40 tapes described earlier.

The 40 reading films were produced in a manner similar to the 40 auding tapes. However, for the films, the rate was determined on a line by line basis. The S value for each line was divided by the Spm for the passage to determine the time in minutes for that line, and then this time was multiplied by 60 to convert to seconds. This time in seconds was then multiplied by 18, i.e., the number of frames per second that the material would be presented. This new value was then the number of motion picture frames of this line that would be shot during the production of the film. For example, if the first line in a passage had 60 characters, its S value would be .60. If the passage was to be presented at 20 Spm, then the time for the first line would be .03 minutes or 1.8 seconds. The 1.8 seconds would then be multiplied by 18 to get 32 frames, i.e., 32 motion picture frames for the first line.

The filming procedure involved the placement of a card (a mask) at the top of the line being shot for the specified number of frames. Thus, in the example above, a card would be placed on top of the first line and then the camera would be shot for 32 frames to produce a 20 Spm rate for the first line. This means that the entire passage was exposed for these 32 frames. Then, the card would be moved down so that it was on top of the first line, leaving all the remaining lines exposed, and then the camera would be shot for the number of frames calculated to provide a 20 Spm rate for the second line. The film was shown using a 16 mm projector that had been modified to present the material at the 18 frames per second rate instead of its normal rate of presentation, 24 frames per second. The effect of the above procedure was to provide an image on a 5' x 6' screen, i.e., the entire passage, that did not move, but each succeeding line on the screen would disappear at the specified rate.

After the 40 films had been produced using the procedures outlined above, it became apparent that this procedure would not provide absolute control over the manipulation of rate. For example, suppose a subject normally read at 15 Spm (i.e., 250 standard words per minute) and the passage was being presented at 30 Spm. Rather than getting halfway through each line before it was covered, the subject could skip to the middle of the passage at the outset and then simply read the second half of the passage at his or her own normal rate. Rather than allow this possible mutilation of the desired effect upon rate, an additional experimental condition was imposed. An apparatus was built which consisted of a piece of cardboard in a frame perpendicularly positioned about four feet in front of the lens of the projector. After the passage was first presented on the screen, this piece of cardboard was raised by the experimenter so that the botton lines of the passage were covered, i.e., masked. Only the first three lines of the passage were exposed. Then, at the moment when the covering mask in the film was lowered over the first line, the experimenter manually lowered

the cardboard mask one line so that there were still only three lines exposed for possible reading. This procedure was repeated so that there were always only three lines exposed at any one point in time, until the last two lines in the passage were reached. This procedure was effective in manipulating reading rate within the constraints desired. However, it should be noted that an individual did have a three line window, i.e., about 30 words, within which there was freedom to skip around and thereby circumvent somewhat the attempt to directly manipulate reading rate.

The rate (R) variable was in Spm. This variable is also considered to reflect the number of thoughts presented per minute (Carver, 1977-78a). In some instances rate is reported in standard length words per minute, Wpm, i.e., Spm multiplied by 16.7.

Accuracy Variable

The primary method of estimating the accuracy with which the subjects had understood the thoughts presented in the reading and auding passages involved subjective judgments by the subjects themselves. This technique has been shown to be reliable and valid for use with prose passages (Carver, 1973), at least when used in conjunction with an objective test. It has been more sensitive to change than objective tests and the curves it has produced are smoother thus indicating more consistancy, i.e., reliability.

Immediately after having been presented a passage, the subjects were asked to estimate the percentage of the total complete thoughts in a passage that they understood. The subjects had been prepared for this estimation procedure during the six practice sessions conducted in the morning. At the outset, the subjects had been informed as follows: (a) "A sentence is a complete thought." (b) "An estimate of understanding could be anywhere between 0 and 100%." (c) "If you think you understood all of the thoughts in the entire passage, you should record a 100% estimate for the passage." (d) "If you understood a few words and phrases,

but you do not think you understood a single complete thought, then you should record a 0% estimate for the passage." (e) "Do not count the exact number of sentences in the passage or the exact number of sentences that you understood."

The numbers 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 were at the top of each test on a passage, and the subjects were instructed to circle the one which provided the best estimate of the percent of complete thoughts in the passage that they thought they understood. These accuracy percents were converted to proportions by dividing by 100 to give an estimated accuracy $(\widetilde{A})$, i.e., the estimated number of thoughts or sentences understood per the number of thoughts or sentences presented.

The test score data will also be considered as indicators of the accuracy of understanding. However, these test scores cannot be considered as direct estimates of $\widetilde{A}$ unless they are specially scaled to reflect the proportion of thoughts understood.

Efficiency Variable

Estimated efficiency $(\widetilde{E})$ is in terms of the estimated number of thoughts or sentences understood, or stored, per minute. For example, if an individual was presented a passage at 12 Spm and the individual estimated that 30% of the thoughts had been understood, i.e., $\widetilde{A} = .30$, then efficiency would be calculated by multiplying .30 by 12 to give an $\widetilde{E}$ value of 3.6 thoughts, or sentences comprehended per minute.

Practice Sessions

As mentioned earlier there were six practice sessions conducted in the morning, without tests. Each session consisted of the presentation of 10 passages. The practice passages used for the first 10 groups were the same experimental passages as were used for the second 10 groups, i.e., the Grade 5 and 14 passages. The practice passages used for the second 10 groups were the same experimental passages as were used for the first ten groups, i.e., the Grade 8 and 11 passages.

Approximately the same design was used in the six morning practice sessions as was used in the four afternoon experimental sessions. However, the passages used in the first and second practice sessions in the morning were exactly the same passages used in the following third and fourth sessions. During the first two sessions, the passages were presented in printed form, i.e., on a sheet of paper, whereas the same passages were presented again on tape or film in sessions three and four. For sessions five and six, the conditions were exactly the same as in the experimental sessions given in the afternoon except there were no tests.

In these morning practice sessions, the subjects were simply asked to record their percent understanding ratings immediately after they were presented a passage. These six sessions provided the subjects with a great deal of procedural practice. They had practice in rating passages for understanding 60 times. They had practice in reading by film at all ten of the rates, twice. They also had practice in auding compressed speech material at all ten of the rates, twice.

The morning design for practice was slightly different from the afternoon design in that the order of the four films and tapes alternated, e.g., one of the six sessions consisted of Read Text, Read Text, Read Film, Listen to Tape, Read Film, Listen to Tape, and the other type consisted of Read Text, Read Text, Listen to Tape, Read Film, Listen to Tape, Read Film.

Data Analysis

For the accuracy estimates, $\bar{A}$, there were either four, five, or six subjects for each passage under each rate condition. The mean of these 4-6 values was calculated, and then the mean of all ten passages at each rate was calculated to give a single accuracy estimate for each of the 10 rates for each of the four difficulty levels. For the two types of test scores, the analysis was the same

except that there were only 2-3 subjects, instead of 4-6, in the original mean. A correction for guessing formula was used on both the RS-test, rights minus 1/4 wrongs, and the P-test, rights minus 1/3 wrongs.

There was a second type of analysis which was similar to the above except the subjects were first divided into ability levels 4 and 5, using scores on the National Reading Standards Test, and the median accuracy, $\tilde{A}$, at each rate and at each level of ability was determined. There were too few subjects at ability levels 3 and 6 to provide reliable data so the data from those subjects were not included in this second type of analysis.

Results and Discussion

The first results to be presented will be those relevant to the measurement of accuracy. If all of the three methods of estimating the accuracy of understanding or comprehending sentences did not produce approximately the same results, then the remainder of the results would have to be presented for each particular technique used to estimate accuracy.

For each of the four levels of material difficulty (Grades 5, 8, 11, and 14), for each of the three methods of estimating accuracy (understanding judgments, RS-tests, P-tests), and for the reading mode as well as the listening mode there is one mean for each of the ten presentation rates (5, 10, 12.5, 15, 17.5, 20, 22.5, 25, 30, ∞ Spm). In order to have a quantitative measure of the degree to which each of the two objective tests were reflecting the same functional relationship with rate as the understanding judgments, the correlations between the objective test score means and the understanding judgment means were computed for the sets of 20 means (10 means for reading plus 10 means for auding) at each of the four difficulty levels. These eight correlations are presented in Table 2. For example, the correlation between the 20 pairs of means for the understanding judgments and the RS-test for Grade 5 material was .86. The correlation for understanding judgments and the P-test at Grade 5 difficulty was .98.

Insert Table 2 about here

It may be noted in Table 2 that all eight of the correlation coefficients were greater than .85 and that the P-test scores tended to correlate higher with the understanding judgments than did the RS-test scores. The size of these correlations suggests that there is not an important difference between

the resulting functional relationship between accuracy and rate using understanding judgments as compared to either of the two objective tests.

Before continuing with the primary data analyses, certain comparisons will be made between the RS-test and the P-test. It has been noted that the P-test tended to correlate slightly higher with the understanding judgments than did the RS-test. This result may be interpreted as suggesting that the P-test, in this study, provided slightly more reliable results than the RS-test.

Another technique for comparing the validity of the objective tests has been developed and used by Carver (1974b, 1975b). This technique compares the amount of gain in percentage points on the objective test per the amount of gain in percentage points in the understanding judgments. This ratio has been termed the efficiency ratio. A zero efficiency ratio would mean zero validity because test scores would be increasing a zero amount for any increase in understanding as reflected by the understanding judgments. A 1.00 efficiency ratio would reflect perfect validity since each percentage point increase in understanding judgments would be associated with a corresponding percentage point increase in test scores. Table 2 also contains the efficiency ratios for the relationship corresponding to the correlations already discussed. These efficiency ratios represented the slope of the regression of the objective test upon the understanding judgments. Notice that these validity estimates tend to be higher for the P-test as compared to the RS-test. The mean of the four P-test values, .61, is almost twice as large as the mean of the four RS-test values, .36. The .36 value is well within the range of RS-test values found in prior research (see Carver, 1974b; Carver, 1975b); the expected validity of the RS-test is around .35 to .45. These findings replicate prior research in that other types of P-tests have tended to have higher efficiency ratios than the RS-test. However,

In the earlier research, the P-test efficiency ratios tended to be more erratic suggesting variability in the competency of the item writer.

Generalizing these results beyond these experiments is a complex matter. Similar results using the RS-test could be expected from other researchers since the RS-test is developed in a completely objective manner. However, the P-test results are likely to be specific to this experimenter. Other researchers may do a better job of developing P-test questions or other researchers may do considerably less well in developing such questions. Until there is better standardization in developing the P-test type of test items, there is likely to be great variability in the results.

It appears that in the present experiment, the P-test might be more valid in the sense that it is more sensitive to changes in understanding than the RS-test. However, both tests appear to reflect the same functional relationship between accuracy and rate as does the understanding judgments since the correlations with the understanding judgments were so high. Rather than simply report the size of these correlations, however, it would be better if a graphic presentation of the relationship was made. Such a presentation is contained in Fig. 1.

Insert Fig. 1 about here

Fig. 1 contains eight graphs, a reading and auding graph for each of the four difficulty levels. Each of the eight graphs presents the relationship between estimated accuracy, $\tilde{A}$, and estimated rate, $\tilde{R}$. The primary rate variable is sentences per min., Spm, but to facilitate interpretation, rate is also presented in standard length words per min., Wpm. Within each graph there are three curves plotted, one curve for each of the three methods of estimating accuracy. For the understanding judgments, each point on the curve represents

the mean of the percent ratings, converted into proportions by dividing by 100. For the two objectives tests, each point is not simply the mean of the test scores. Rather than present the mean of test scores, a linear transformation of these means was calculated so that the test scores were on the same scale as the understanding judgments. This linear transformation was accomplished by using the same correlations and regression equations that were discussed in connection with Table 2 to predict a mean understanding judgment score for each mean test score. This rescaling procedure has absolutely no effect upon the shape of the curves for the test scores. It simply has the effect of making the units of measurement equal for all three methods of estimating accuracy. It is similar to converting scores into z-scores or T-scores since the newly scaled scores have exactly the same mean and standard deviation as the understanding judgments.

It may be noted in Fig. 1 that the two objective test curves, indicated by dotted lines, are almost perfectly coincident with the subjective test curve, indicated by a solid line, in all eight of the graphs. There appears to be no evidence which would suggest that these three methods are not reflecting the same functional relationship between accuracy and rate. It appears reasonable to suggest that the particular method used for estimating accuracy does not interact with the functional relationship, at least it did not in the present experiment. Thus, it appears that rauding theory can be tested using multiple methods of estimating the accuracy of comprehending thoughts.

Before leaving the data in Fig. 1, it should be noted that these data would be interpreted as supporting Jester and Travers (1966) in their contention that the relationship between rate and comprehension was linear. Indeed, these data present the same general relationship as was also presented by Jester and Travers. However, the approximately linear function is only superficial, as will be noted in later analyses involving 1/E versus 1/R.

The next results to be considered will focus primarily upon the question of whether there is an optimal rate, and secondly, the results will provide further information relevant to the degree to which the subjective judgments are consonant with the objective test data. Fig. 2 contains eight graphs presenting the relationship between efficiency, $\tilde{E}$, and rate, $\tilde{R}$, for both reading and auding at the four levels of difficulty. For the understanding judgments, the $\tilde{E}$ values were derived by simply multiplying the $\tilde{A}$ values given in Fig. 1 by their corresponding $\tilde{R}$ values. The data for the objective test curves in Fig. 2 were derived by calculating the mean of the two test means at each rate in Fig. 1, and then multiplying this average $\tilde{A}$ value by its corresponding $\tilde{R}$ value. Therefore, the objective test data in Fig. 2 were derived from the data given in Fig. 1 exactly the same way as was the understanding judgments except the curves for the two objective tests were first collapsed into one objective test curve.

Insert Fig. 2 about here

Notice, in Fig. 2, that there appears to be no question but that there is an optimal rate since efficiency increases as rate increases initially, but then efficiency tends to decrease at the high rates. The decline in efficiency is more pronounced for auding than it is for reading. This is probably because compressed speech tends to distort the words at the high rates, i.e., a loss in intelligibility, whereas in the reading situation the subject could skip around within the three line window described earlier and actually read some of the words at a rate somewhat slower than the purported rate. An inspection of the reading data in Fig. 2 suggests that a maximum efficiency exists; therefore, an optimal rate exists as well. The only questionable aspect of the reading data involves the objective test results for Grade 5 difficulty. These data suggest an optimal,

rate at around 17.5 Spm, except at 30 Spm efficiency increases again. Since there is no question about the existance of an optimum for the other seven curves depicting the relationship between efficiency and rate of reading, it would seem most likely that this single data point is simply unreliable and inaccurate. In summary, these data may be interpreted as supporting that part of rauding theory which holds that there is an optimal rate of comprehension which is associated with a maximum efficiency of comprehension.

These data in Fig. 2 also provide additional information relevant to the consistency between the subjective judgments of the subjects and objective test scores. The curves of the understanding judgments and the objective test scores are almost perfectly coincident. Again, these data indicate that there is an optimal efficiency and a corresponding optimal rate regardless of whether the efficiency estimate was derived from objective test scores or from subjective judgments.

The question of whether the optimal reading rate is different from the optimal auding rate was investigated by averaging the highest points in Fig. 2. The three highest points on the curve for the understanding judgments and the three highest points on the curve for the objective tests were combined and an average value for these six data points was calculated. These data are presented in Fig. 3. For example, in Fig. 2, the mean of the six highest efficiency values from the Grade 5, Level 2, reading data was 13.1 Spm and this value has been plotted in Fig. 3 for Grade 5. The mean of the six reading rates that were associated with the six highest reading efficiency values was 19.2 Spm and this value has also been plotted in Fig. 3 for Grade 5. It may be noted that the optimal efficiency values for reading were almost perfectly equal to that of auding at each level of difficulty, i.e., the optimal reading efficiency curve and the optimal auding efficiency curve were almost perfectly coincident. These data suggest

that there is nothing inherent in the reading mode which makes it more efficient for understanding sentences in passages than the auding mode. The superiority of reading comes from the ability of the individual to read at 300 Wpm under typical reading conditions while the individual normally cannot aud at 300 Wpm because auding rate is typically around 150-175 Wpm and beyond the individual's control. Reading is also superior to auding at the high rates, e.g. 417 and 500 Wpm, but this is probably due to a deterioration of the stimulus in auding (called loss of intelligibility). When auding rate is mechanically manipulated so that it equals the rate at which reading is most efficient, there is no significant difference between the efficiency of reading comprehension and listening comprehension.

Insert Fig. 3 about here

The rates that are associated with the optimal auding efficiency are almost perfectly constant, i.e., around 18 Spm, across the four levels of difficulty. The corresponding rates for the reading data are not as consistent since the optimal rate at Grade 8 is higher than the other three values. However, for the other three grades the optimal rate was around 18 Spm or 300 Wpm. These data suggest that the optimal reading rate might be slightly higher than the optimal listening rate, but because of the slightly erratic nature of one data point, it does not seem that any strong generalizations are justified from this data analysis. Other data to be presented later from the second type of data analysis provides more consistent results. (It should be noted that for the data presented in Fig. 3, the rise in the reading efficiency curve at 30 Spm for Grade 5, mentioned earlier, was considered to be unreliable and was not included as one of the six highest values.)

The next data to be presented will be directly relevant to the question of whether the relationship between accuracy and rate is a hyperbolic function

represented by Equation 3. As mentioned earlier, if there is a linear relationship between $1/\tilde{E}$ and $1/\tilde{R}$, as indicated by Equation 2, then this will provide evidence for a hyperbolic function instead of a linear function relating A and R. Fig. 4 contains two graphs relating $1/\tilde{E}$ to $1/\tilde{R}$, one graph for reading and one for auding. The values in Fig. 4 were derived from the values in Fig. 2. First, the two curves in each graph in Fig. 2 were combined by calculating the mean of the set of two means at each rate, i.e., the mean of the understanding judgment mean and the objective test mean. This combination of all the estimates of accuracy should provide the most reliable data. These mean efficiency values, $\tilde{E}$, at each rate were used to calculate the estimate of the time per sentence understood, $1/\tilde{E}$, and these values were plotted against the time per sentence presented, $1/\tilde{R}$. It may be noted in Fig. 4 that when the amount of time exceeded about .05 minutes per sentence (20 Spm or 333 Wpm), the relationship between 1/E and 1/R was almost perfectly linear in both the reading and auding modes. For the eight curves in Fig. 4, a straight line was fit to the four data points that had the following $1/\tilde{R}$ values: .067, .10, .08, and .02. This straight line was fit to the data for each of the four difficulty levels in both the reading and auding modes. Table 3 contains the correlations coefficients for the eight linear relationships. Notice that a straight line seems to fit these relationships extremely well. Two of the correlations were .99, and the other six correlations were perfect, 1.00, when rounded to two decimal places. Also in Table 3 are the slopes of the regression lines. Notice that the slopes of the lines are all very close to unity. They range from 0.9 to 1.1 and the mean of all eight is 1.0. These data indicate that the general theoretical equations, Equations 2 and 3, fit these data very well, i.e., a hyperbolic function relates $\tilde{A}$ to $\tilde{R}$ with only the y-intercept, I, varying.

Insert Table about here

These results may be interpreted as indicating that for college students who are presented materials at Grades 5, 8, 11, and 14 difficulty, the relationship between accuracy and rate of comprehending sentences is best represented by the hyperbolic function of Equation 2 when the rate is below the optimal rate for both reading and auding. These data indicate that the approximately linear relationship between rate and comprehension found by Jester and Travers (1966) and replicated in Fig. 1 is only superficial. Stated differently, these data support that part of rauding theory which holds that for rates below the rauding rate, i.e., below the optimal rate, the relationship between the time per sentence comprehended, 1/E, and the time per sentence presented, 1/R, is the linear function represented by Equation 2.

Insert Fig. 4 about here

The next data analysis has two purposes. One purpose is to compare reading with auding. The other purpose is to compare the obtained results with the results that may be predicted from rauding theory. Fig. 5 contains the same data as are contained in Fig. 4 except there is a graph for each of the four difficulty levels. Reading may be compared with auding on each of the four graphs. It may be noted that there are no major differences between the 1/E data points for reading and auding at each value of 1/R. These data may be interpreted as supporting the contention of rauding theory that there are no major comprehension differences between the two modes.

Insert Fig. 5 about here

The approximately 300 word per min. optimal rate of reading replicates the Jester and Travers results but they found a 200 word per min. optimal rate for

auding whereas a 300 Wpm value was found for auding in the present research. The lower optimal rate found by Jester and Travers possibly reflects the fact that they gave their subjects no prior practice in listening to time compressed speech, a limitation which they alluded to themselves.

In Fig. 5 it may be noted that the predicted curves fit the data quite well, especially the predicted linear relationship between $1/\tilde{E}$ and $1/\tilde{R}$ beyond the threshold or optimal rate. The predicted optimal rate is indicated in Fig. 5 by the vertical line. The optimal rate, called the rauding rate, R_r, for the data in all four graphs was estimated to be 18.3 Spm, i.e., $\tilde{R}_r$=18.3 and $1/\tilde{R}_r$=.055. Thus, the vertical line representing the optimal rate lies on a $1/\tilde{R}$ value of .055. The 18.3 Spm value was derived from the median grade ability of the 102 Ss in the experiment, i.e., the median Grade Ability score, G_a, on the NRS test for the 102 S_s was 13.5. Using the table provided in the presentation of rauding theory (Carver, 1977-78), the estimated rauding rate for this score was determined to be 305 Wpm, and the R_r value is equal to 18.3 Spm. Then, the same table was used to estimate the four rauding accuracies ($\tilde{A}_r$) associated with the differences between the ability level and difficulty level ($\tilde{G}_a - \tilde{G}_d$). Rauding accuracy, A_r, is simply the accuracy, A, that accompanies the rauding rate. The four $\tilde{A}_r$ values associated with the $\tilde{G}_a - \tilde{G}_d$ values were .91, .85, .72, and .48, respectively. Each of these four values were then substituted into an equation for I given in rauding theory,

$$I = \frac{1}{R_r}\left(\frac{1}{A_r} - 1\right), \qquad (4)$$

together with the constant value of rauding rate, R_r = 18.3. The resulting I values were .005, .010, .021, and .059, respectively. The latter I values are the $1/\tilde{E}$ intercepts of the four prediction curves and the actual data points provides strong support both for rauding theory and for the suggested empirical procedures for estimating A_r and R_r.

Optimal Rate

The preceding data have represented the mean or average of all the 102 college students who participated in the research. The next analysis will focus upon the differences between individuals at Level 4 ability (Grade abilities 10, 11, and 12) and Level 5 ability (Grade abilities 13, 14, and 15). In this analysis, it will be possible to be more precise when investigating the interaction between the level of ability, L_a, and the level of difficulty, L_d. Since the preceding analyses found no important differences between the estimates of accuracy that came from understanding judgments as compared to the objective test scores, the following analyses were simplified by analyzing only the understanding judgments. As explained earlier, the median accuracy estimates were determined for the individuals at each of the two levels of ability represented, i.e., $L_a = 4$ and $L_a = 5$, and all subsequent data were derived from these medians. One disadvantage of this analysis was that the extensive design controls that applied to the earlier analyses no longer apply. However, since there were a large number of individuals in each of the two groups, and since the passage order differences were not expected to produce drastic differences in the results, this disadvantage should be minor.

Fig. 6 contains four graphs of efficiency versus rate--Level 4 reading, Level 4 auding, Level 5 reading, and Level 5 auding. Each graph contains a curve for each of the four grade difficulties. It may be noted that these curves present relationships which are very similar to those presented in Fig. 2, i.e., there is an optimal efficiency since efficiency is a maximum somewhere around 17.5 to 20 Spm.

Insert Fig. 6 about here

The data in Fig. 6 were subjected to the same type of analysis as was

described earlier, in connection with Fig. 3, for getting average optimals. The three highest efficiencies for each of the four grade abilities in each of the four graphs were averaged, i.e., a mean was calculated. The three rates that were associated with the three highest efficiencies were also averaged. These values are presented in the two graphs of Fig. 7.

Insert Fig. 7 about here

In general, the Level 5 ability results replicate the Level 4 ability results. For both Level 4 and 5 ability, the reading efficiencies are higher than the auding efficiencies but the difference is extremely slight, i.e., about 1 to 2 Spm, or about a 10% absolute difference. There appears to be a definite drop in efficiency at Grade 14. For the optimal rates, the results are more consistent than was presented earlier in Fig. 3. There seems to be a fairly stable optimal rate across all four difficulty levels for both reading and auding at each ability level. Furthermore, this optimal rate seems to be consistently higher for reading than for auding.

Table 4 contains an average (mean) of the four optimal rates given in Fig. 7. For example, the mean of the four optimal rates of reading for the Level 4 individuals was 18.9 Spm, i.e., averaging across the four difficulty levels in Fig. 7. The average rate values are also presented in Wpm. Notice that the Level 5 values are about 1-2 Spm higher than the Level 4 values. For the reading versus auding comparison, reading is about 2-3 Spm higher than auding.

Insert Table 4 about here

Optimal Rate

Also contained in Table 4 are predicted values for the optimal rate. These values come from the same rauding theory table (Carver, 1977-78a) noted earlier. The score these individuals made on the standardized reading test was used to estimate their optimal rate, i.e., R_r. Notice that the predicted optimal rate values are comparable to the actual optimal rates found for reading and auding. For example, the predicted optimal rate for the Level 5 individuals was 316 Wpm and the average optimal rate for reading was 347 Wpm and the average optimal rate for auding was 295 Wpm. It appears that the empirical predictions made from rauding theory regarding rauding rate or optimal rate are relatively accurate.

These rate data suggest support for the theory that both Level 4 and Level 5 individuals have an optimal rate at which they can understand the sentences in passages, and that this optimal rate remains relatively stable or constant across difficulty levels and across the two modes of presentation. The optimal reading rate was only slightly higher than the optimal auding rate, i.e., about 1-2 Spm on the average.

The relationship between accuracy and difficulty for the two levels of ability during both reading and listening was also investigated. The accuracy data, $\widetilde{A}$, that corresponds to the efficiency, $\widetilde{E}$, and rate, $\widetilde{R}$, data in Fig. 7 are plotted in Fig. 8, i.e., $\widetilde{A}=\widetilde{E}/\widetilde{R}$. Notice that accuracy values are around .70 to .80 for Levels 2, 3, and 4 difficulty for both Level 4 and Level 5 ability in both reading and auding. For Level 5 difficulty, accuracy drops off to about .50 in all four curves.

Insert Fig. 8 about here

The values in Fig. 8 may be referred to as rauding accuracy, A_r, values since they represent the accuracy that accompanied the rauding rate or the optimal rate. There does not seem to be a tendency for the rauding accuracies

values for reading to exceed those for auding. This is in accordance with rauding theory. However, there also does not seem to be a tendency for the Level 5 accuracies to exceed the Level 4 accuracies. This equivalency was not predicted by rauding theory. An interpretation of this inconsistency will be facilitated by f rther data analyses. The mean of the four accuracy values at each difficulty level was calculated and the results were .78, .77, .74, and .48 for difficulty levels 2-5, respectively. These values are roughly equal to the values that would be predicted for Level 5 individuals--.92, .87, .76, .53. These values tend to be higher than predicted for Level 4 individuals--.87, .76, .53, .22. It appears that these college level students who tested at Level 4 (high school level) tend to read at a rate similar to high school students but their accuracy of comprehension tends to be more like their college student peers. Rather than interpreting this inconsistency as reflecting upon rauding theory, it will be interpreted as reflecting upon the empirical measure used to estimate A_r and R_r. It appears that the score on NRS test provides a relatively accurate estimate of rauding rate at the higher levels but it does not provide a highly accurate estimate of rauding accuracy. In this case, the status of the reader, i.e., college students, would provide a more accurate predictor of what material they could accurately understand than the reading test from whence the ability levels were designated. The unpredicted equivalency between the rauding accuracies of the Level 4 and Level 5 individuals will be interpreted as indicating that better measures are needed for estimating A_r at the college level.

These data analyses for Level 4 and Level 5 are consistent with most of the interpretations made from the earlier data analyses. It appears that Level 4 and Level 5 individuals have an optimal rate of reading and auding. Rates either above or below the optimal rate result in lower efficiency of comprehending sentences in passages. In this study there was a definite tendency for reading to have a higher optimal rate than auding. However, rather than concluding that

this is evidence against rauding theory which holds that there is a common optimal rate that is centrally determined, just the opposite is advocated. A difference of 2 Spm between reading and auding for Level 5 individuals who have an optimal reading rate of about 20 Spm is only a 10% difference. A difference of only 10-15% would seem to be too small a difference to reject the hypothesis of no difference. A difference this small could easily be an artifact of this specific experiment. For example, it could be that a 10% increase in the optimal auding rate could be affected by using a better compressed speech machine or by having the recording announcer read the original passage at a higher rate so that less compression would be needed (and, therefore, less loss of intelligibility of the words) to get the higher rates. Or, it could be that the higher optimal rate of reading resulted simply from the fact that the readers always had three lines visible from whence they could read some sentences at their own rate and be less affected by the rate at which the material was supposedly presented. It may be that presenting only one line at a time for reading, or one sentence at a time, or even one word at a time would have decreased the optimal rate by 10%. It does not seem reasonable to reject the theory of a centrally determined optimal rate when the difference was no larger than 10-15%, a difference well within the expected accuracy error of the experimental techniques used.

Summary and Conclusions

The empirical results of this research provide strong support for the existence of an optimal rate of reading prose and an optimal rate of auding prose. The optimal rate of reading prose tended to be approximately equal to the optimal rate of auding prose. The efficiency of comprehension for this group of college students tended to peak at an optimal rate of approximately 300 words per min.

for reading and auding prose material at grade difficulties of 5, 8, 11, and 14. The efficiencies associated with the optimal rates also tended to be equal across both modes. In summary, the data indicate that there is a constant optimal rate across modes and across prose difficulty levels.

The data alluded to above provide support for rauding theory, as outlined at the outset. Also supporting rauding theory was: (a) the consistent linear relationship found between $1/\widetilde{E}$ and $1/\widetilde{R}$ which had a slope of 1.0, and the close correspondence between the 1/E values predicted from rauding theory and the empirical values of $1/\widetilde{E}$. This latter result provides automatic support for the purported hyperbolic relationship between the accuracy of comprehension and rate, Equation 3. In turn, these data provide evidence against a linear relationship, suggested by Jester and Travers (1966), between comprehension and rate .

These data do support rauding theory, but they do <u>not</u> appear to be in accordance with conventional ideas about the characterists of a good reader as seen from the hypothesis testing view. Good readers, i.e., flexible, mature readers, are supposed to be able to hypothesize what lies ahead and there upon adjust their rate so as to match changes in the difficulty or redundancy of the material. Seemingly, this would mean that as the difficulty of the material decreased, then individuals would increase their rate so as to increase their reading efficiency. The present data suggest that approximately 300 Wpm is the most efficient rate for typical college students when they read college level material as well as when they read grade 5 materials. Thus, it would not seem appropriate for good readers to adjust or change their rate as material decreases in difficulty. Indeed, the research referrenced in the introduction regarding the constancy of reading rate indicates that individuals do not adjust their rate to such changes in material difficulty, so it seems that individuals are probably maximizing their efficiency of reading prose by keeping their rate constant at the optimal rate.

It does appear reasonable to theorize that individuals have a rather constant rate of reading prose, called the rauding rate. At this rate individuals are able to maximize their effiency. When individuals increase their rate above the rauding rate, then it appears reasonable to consider the accompanying cognitive processes as qualitatively different. This view is in opposition to that of Hochberg and Brooks (1970), for example, who have contended that "...'skimming' and 'reading' represent points on a continuum of intention; they are not basically different activities [p. 306]." Rauding theory and its supporting data suggest that future theory regarding the detailed cognitive processes during reading should make a hard distinction between typical prose reading, at rates around 300 words per min., and skimming, at rates around 1200 words per min.

With respect to the equivalency of reading and auding prose, there appears to be little evidence that individuals can accurately comprehend the sentences in passages during reading any faster than they could if the material was presented auditorily at the same rate. There was little or no difference between reading and listening comprehension at any rate, especially the optimal rate. There was evidence that the optimal rate of reading might have been around 30-40 Wpm higher than for auding but this slight difference could easily be accounted for by the experimental techniques used to manipulate reading and auding rate. This means that the uniqueness of reading comprehension as compared to listening comprehension is very small. This conclusion is in keeping with that of Larsen and Feder (1940) who compared reading and auding and concluded that their data supported "...the conclusions of previous studies that comprehension is largely a centrally-determined function operating independently of the mode of presentation of the material [p. 251]." More recently a similar type of conclusion was drawn by Rosenberg and Simon (1977) who contended that "...subjects integrate information across modalities into a single underlying semntic representation [p. 293]."

it appears that individuals learn to read with a certain amount of proficiency which equals but does not exceed that of their initial modality, listening. Indeed, one of the major goals of good reading might be regarded as being reached when individuals learn to read as efficiently as they can listen.

The empirical results directly supporting Equation 2 deserve further discussion. Since the linear relationship of Equation 2 has a slope of unity, this means that if 1/R is increased by a certain amount, 1/E will therefore be increased by exactly the same amount. The units of 1/R and 1/E are both "time per thought." Therefore, when the amount of time allowed per each thought presented is increased by 2 seconds, for example, then the amount of time taken to comprehend each thought is also increased by 2 seconds. This relationship may be referred to in rauding theory as an empirical law. The verbal description of this law is as follows: An increment in the amount of time per thought presented results in an equal increment in the amount of time per thought comprehended. This law only applies when the individual has more than ample time to finish reading a passage, i.e., when the time allowed for reading exceeds the amount of time required to read the material at the optimal rate. This empirical law of rauding seems to hold for college students who are reading material with a wide range of grade difficulties, i.e., Grade 5 through Grade 14. This law is especially useful for predicting comprehension in situations where individuals have been given more than enough time to read a passage once at their normal rate. Theory is lacking regarding "why" the law works as it does.

For the results summarized above, three drastically different measures of the accuracy of comprehension produced essentially the same results. Having individuals rate their percent understanding of passages produced results that were comparable to multiple-choice paraphrase questions on each sentence and to

reading-storage test questions. Thus, the above evidence for rauding theory appears to be generalizable beyond any one particular measurement technique.

In summary, the evidence presented in this article has been interpreted as supporting the existence of an optimal rate at which sentences in prose passages can be comprehended. This optimal rate appears to be approximately the same rate in both the reading mode and the auding mode. The optimal rate seems to increase as the reading ability of the individual increases. But, for any particular ability level, the optimal rate appears to be constant across different difficulty levels. When the rate of presentation of the material is less than the optimal rate, the relationship between rate and accuracy seems to be the hyperbolic function represented earlier by Equation 3. Since all of these data regarding optimal rate were primarily predicted from properties of the rauding rate construct, R_r, it appears that rauding rate, R_r, is a viable theoretical construct. Furthermore, there appears to be enough supporting data for rauding theory itself to justify its use when designing both basic and applied research in reading. Theories about the cognitive processes involved in reading need to account for (a) the optimal rate, (b) the equivalencies of reading and auding, and (c) the constancy of the optimal rate across varying difficulties of prose materials.

References

Bormuth, John R. Development of readability analyses. U. S. Office of Education Final Report, Proj. No. 70052, Contract No. OEC-3-7-070052-0326, University of Chicago, March, 1969.

Carroll, John B. On learning from being told. Educational Psychologist, 1968, 5 (2), 1, 5-10.

Carver, Ronald P. Evidence for the invalidity of the Miller-Coleman Readability Scale. Journal of Reading Behavior, 1972, 3, 42-47.

Carver, Ronald P. Understanding, information processing, and learning from prose materials. Journal of Educational Psychology, 1973, 64, 76-84.

Carver, Ronald P. Manual for the Rauding Scale Qualification Test. Kansas City, Mo.: Revrac Publications, 1974. (a)

Carver, Ronald P. Measuring the primary effect of reading: Reading storage technique, understanding judgments, and cloze. Journal of Reading Behavior, 1974, 6, 249-274. (b)

Carver, Ronald P. Revised procedures for developing reading-input materials and reading-storage tests. Journal of Reading Behavior, 1975, 7, 155-172. (a)

Carver, Ronald P. Comparing the reading-storage test to the paraphrase test as measures of the primary effect of reading. Journal of Educational Psychology, 1975, 67, 274-284. (b)

Carver, Ronald P. Measuring prose difficulty using the Rauding Scale. Reading Research Quarterly, 1975-76, 11, 660-685.

Carver, Ronald P. Word length, reading rate, and prose difficulty. Journal of Reading Behavior, 1976, 8, 193-203.

Carver, Ronald P. Technical Manual for the National Reading Standards. Kansas City, Mo.: Revrac Publications, 1977.

Carver, Ronald P. Toward a theory of reading comprehension and ruading. Reading Research Quarterly, 1977-78, 13, 8-63 (a)

Carver, Ronald P. Another look at Rauding Theory. Reading Research Quarterly, 1977-78, 13, 116-132 (b)

Carver, Ronald P. Constancy of reading rate. Unpublished manuscript, 1978. (d)

Goldstein, Harry Reading and listening comprehension at various controlled rates. Doctoral dissertation, Teachers College, Columbia University, 1940.

Haber;. R. N. Visual perception. In L. L. Elliot and J. A. Swets (Eds.) Psychology and the handicapped child. U. S. Department of Health Education and Welfare, DHEW. Pub. No. (OE) 73-05000, 1974.

Hochberg, Julian and Brooks, V. Reading as an intentional behavior. In H. Singer and R. B. Rudell (Eds.) Theoretical models and processes of reading. Newark, Del.: International Reading Association, 1970. /pp. 304-314/

Jester, R. E. & Travers, R.M.W. Comprehension of connected meaningful discourse as a function of rate and mode of presentation. Journal of Educational Research, 1966, 59, 297-302.

Kleiman, Glenn M. Speech recoding in reading. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 323-339.

Larsen, Robert P. and Feder, D. D. Common and differential factors in reading and hearing comprehension. Journal of Educational Psychology, 1940, 31, 241-252.

Miller, Gerald R. and Coleman, Edmund G. The measurement of reading speed and the obligation to generalize to a population of reading materials. Journal of Reading Behavior, 1972, 3, 48-56.

Neisser, Ulric. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

Poulton, E. C. Time for reading and memory. British Journal of Psychology, 1958, 49, 231-245.

Rosenberg, Steven & Simon, Herbert A. Modeling semantic memory: Effects of presenting semantic information in different modalities. Cognitive Psychology, 1977, 9, 293-325.

Smith, F. Understanding reading. New York: Holt, Rinehart, and Winston, 1971.

Sticht, Thomas G., Beck, L. J., Hauke, R.N., Kleiman, G.M., & James, J.H. Auding and reading. Alexandria, Va.: Human Resources Research Organization, 1974.

Thorndyke, Perry W. Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 1977, 9, 77-110.

Footnote

1. This research was funded by the National Institute of Education, Grant No. NE-G-00-3-0043. The author wishes to thank Fran Cohen and Debra Stein for their help in the data collection and data analysis. The preparation of the manuscript was aided by the constructive criticisms of S. Jay Samuels, Robert Fulmer, and students of the author. Reprints may be ordered from the author, Educational Research and Psychology Dept., School of Education, University of Missouri-Kansas City, Kansas City, Mo. 64114.

2. Auding is an infrequently used term. It will be used to refer to situations where individuals are listening to speech in order to understand the thoughts being presented by the speaker (see Sticht, et. al., 1974).

Table 1

Passage, Rate, and Order of Presentation for the Ten Films or Tapes at Grade 5 Difficulty

Order of Presentation

Film or Tape	1	2	3	4	5	6	7	8	9	10
1-5A	5A2/30	5A3/10	5A4/17.5	5A9/20	5A5/15	5A1/22.5	5A8/5	5A6/25	5A0/00	5A7/12.5
2-5B	5B6/10	5B1/20	5B0/30	5B7/22.5	5B9/17.5	5B4/25	5B8/15	5B3/12.5	5B2/5	5B5/00
3-5A	5A3/20	5A5/22.5	5A5/10	5A1/25	5A8/30	5A6/12.5	5A7/17.5	5A2/00	5A4/15	5A0/5
4-5B	5B6/22.5	5B3/25	5B2/20	5B0/12.5	5B9/10	5B8/00	5B1/30	5B4/5	5B5/17.5	5B7/15
5-5A	5A5/25	5A2/12.5	5A3/22.5	5A9/00	5A6/20	5A1/5	5A7/10	5A0/15	5A4/30	5A8/17.5
6-5B	5B4/12.5	5B6/00	5B9/25	5B3/5	5B8/22.5	5B2/15	5B7/20	5B1/17.5	5B0/10	5B5/30
7-5A	5A1/00	5A5/5	5A9/12.5	5A3/15	5A8/25	5A6/17.5	5A0/22.5	5A7/30	5A4/20	5A2/10
8-5B	5B7/5	5B6/15	5B3/00	5B2/17.5	5B8/12.5	5B9/30	5B0/25	5B1/10	5B4/22.5	5B5/20
9-5A	5A1/15	5A3/17.5	5A9/5	5A6/30	5A4/00	5A8/10	5A5/12.5	5A0/20	5A7/25	5A2/22.5
10-5B	5B0/17.5	5B3/30	5B9/15	5B4/10	5B6/5	5B8/20	5B7/00	5B5/22.5	5B1/12.5	5B2/25

Optimal Rate

Table 2

Correlations between each of the Two Objective Tests and the Understanding Judgments, and Their Corresponding Efficiency Ratios, for the Four Levels of Material Difficulty

Grade Difficulty	Level of Difficulty	Correlations		Efficiency Ratios	
		RS-test	P-test	RS-test	P-test
5	2	.86	.98	.43	.58
8	3	.89	.94	.31	.55
11	4	.93	.96	.26	.66
14	5	.95	.94	.43	.67

Table 3

Correlations and Slopes Associated with the Eight Regression Lines Which Relate $1/\bar{E}$ to $1/\bar{R}$ for Reading and Auding at Four Levels of Difficulty

Grade	Level	Correlations		Slopes	
		Reading	Auding	Reading	Auding
5	2	1.00	1.00	1.1	0.9
8	3	1.00	1.00	1.0	0.9
11	4	.99	1.00	1.0	0.9
14	5	.99	1.00	1.1	1.0

Table 4

Average Optimal Rate of Reading and Auding and Predicted Optimal Rate Values

	Average Optimal Rate				Predicted Optimal Rate	
	Reading		Auding			
Level	Spm	Wpm	Spm	Wpm	Spm	Wpm
4	18.9	315	16.8	280	15.4	256
5	20.8	347	17.7	295	19.0	316

Figures

Fig. 1 Estimated accuracy of comprehension, $\widetilde{A}$, as a function of estimated rate, $\widetilde{R}$, when accuracy is estimated using three different methods--understanding judgments, reading-storage tests (RS-tests) and paraphrase tests (P-tests). There are eight graphs--for difficulty levels 2-5, i.e., L_d=2, L_d=3, L_d=4, and L_d=5, and for both reading and auding at each difficulty level.

Fig. 2 Estimated efficiency of comprehension, $\widetilde{E}$, as a function of estimated rate, $\widetilde{R}$, when accuracy is estimated two ways--understanding judgments, and the two objective tests combined by using the mean of the RS-test and P-test. There are eight graphs--for difficulty levels 2-5, i.e., L_d=2, L_d=3, L_d=4, and L_d=5, and for both reading and auding at each difficulty level.

Fig. 3 Estimated optimal efficiencies, in sentences per minute, as a function of difficulty level, L_d, and the rates, also in sentences per minute, that were associated with the optimal efficiencies at each difficulty level under both the reading and auding conditions. Accuracy was estimated using a combination of both the understanding judgments and objective tests.

Fig. 4 Estimated number of minutes required for each sentence comprehended, $1/\widetilde{E}$, as a function of the estimated number of minutes allowed for each sentence presented, $1/\widetilde{R}$, at the four different difficulty levels when accuracy was estimated using the understanding judgments and the objective tests combined. There are two graphs--one graph presents these data for the reading condition and another graph presents these data for the auding condition.

Optimal Rate

Fig. 5 Estimated number of minutes required for each sentence comprehended, $1/\widetilde{E}$, as a function of the estimated number of minutes allowed for each sentence presented, $1/\widetilde{R}$, under both reading and auding conditions when efficiency was estimated using the understanding judgments and the objective tests combined. There is a graph for each of the four difficulty levels. From rauding theory the optimal rate, $\widetilde{R}_r$, was predicted and is designated by a vertical line on each graph at $1/R_r$. Also from rauding theory, the values of $1/\widetilde{E}$ for each value of $1/\widetilde{R}$ below the optimal rate were predicted and these values are designated by a diagonal line on each graph. The arrows pointing upon each graph above a data point for auding indicates that there were one or more additional data points for auding that had such high $1/\widetilde{E}$ values they could not be included on the graph.

Fig. 6 Estimated efficiency of comprehension, $\widetilde{E}$, as a function of estimated rate, $\widetilde{R}$, for each difficulty level when accuracy was estimated using the understanding judgments. There are four graphs--for ability levels 4 and 5, i.e., L_a=4 and L_a=5, and for both reading and auding at each ability level.

Fig. 7 Estimated optimal efficiencies, in sentences per minute, as a function of difficulty level, L_d, and the rates, also in sentences per minute, that were associated with the optimal efficiencies at each difficulty level, L_d, when accuracy was estimated using the understanding judgments. There are two graphs--for level 4 ability, L_a=4, and Level 5 ability, L_a=5.

Fig. 8 Estimated accuracy, $\widetilde{A}$, that was associated with the optimal efficiencies, as a function of difficulty level, L_d, under the reading and auding conditions for each of the two ability levels, 4 and 5, when accuracy was estimated using understanding judgments.

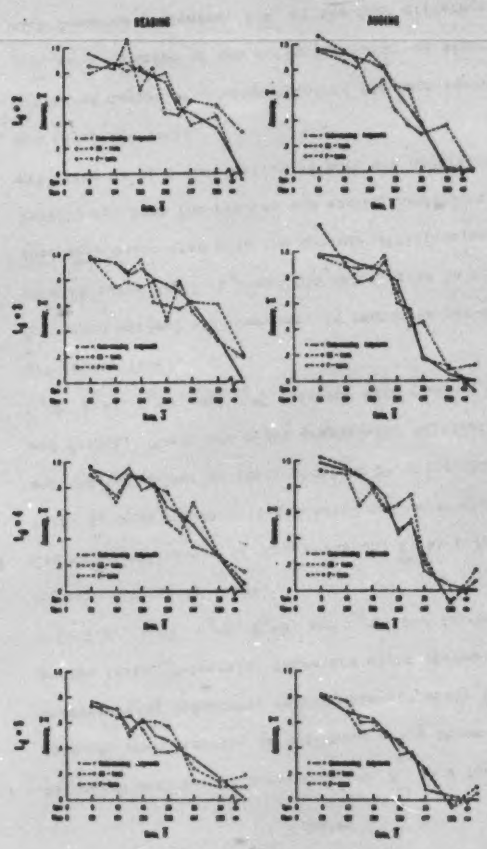


FIG 1

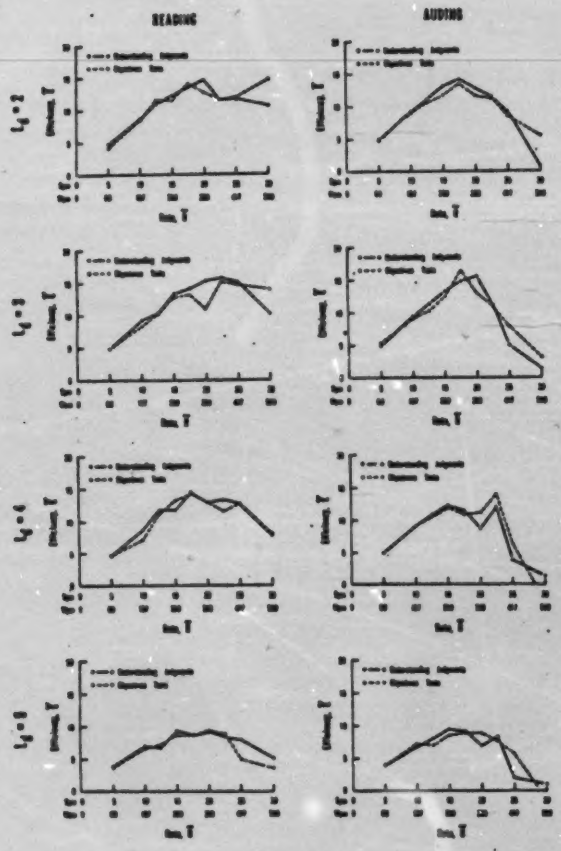


FIG. 2

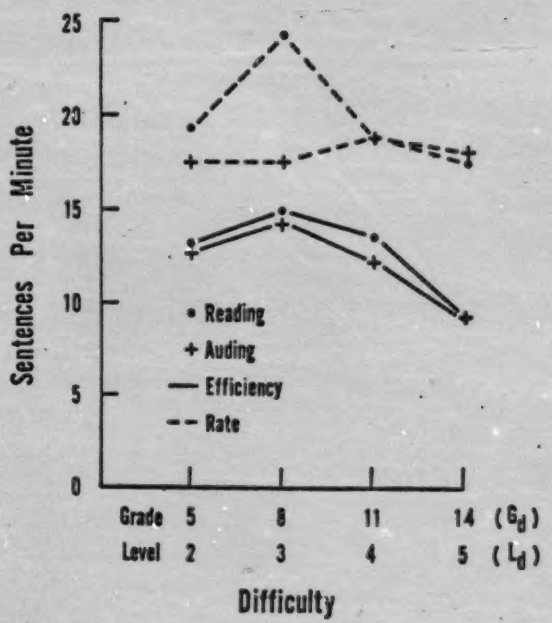

Sentences Per Minute
25
20
15
10
5
0
• Reading
+ Auding
— Efficiency
-- Rate
Grade 5 8 11 14 (G_d)
Level 2 3 4 5 (L_d)
Difficulty

FIG. 3

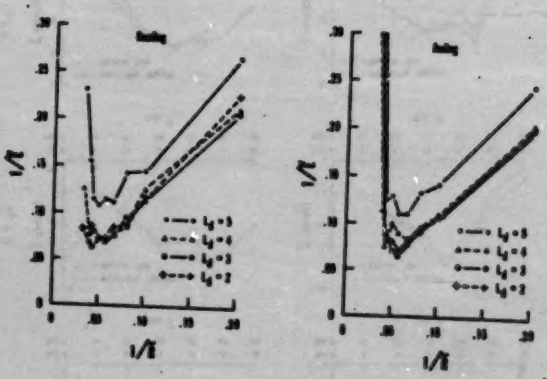


FIG. 4

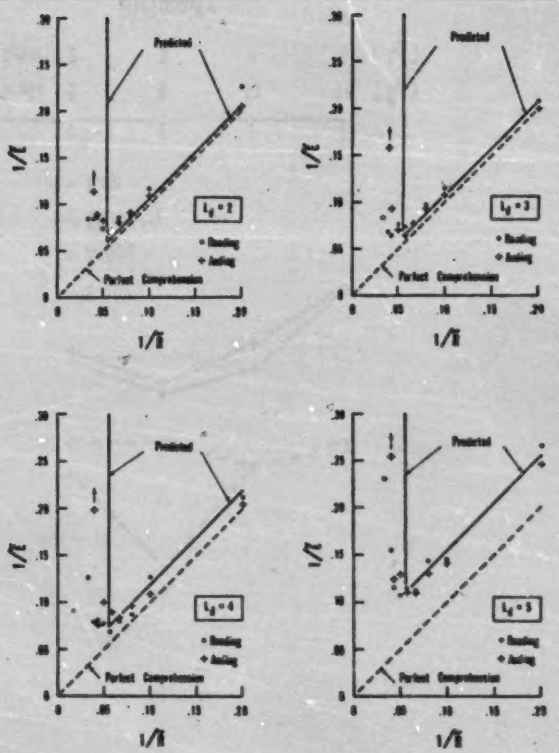


FIG. 5

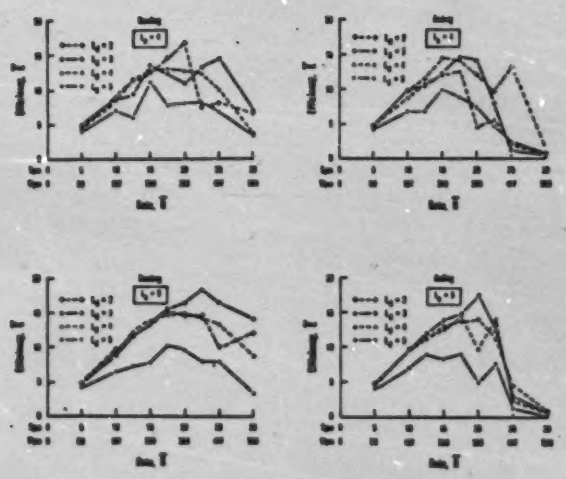


FIG. 6

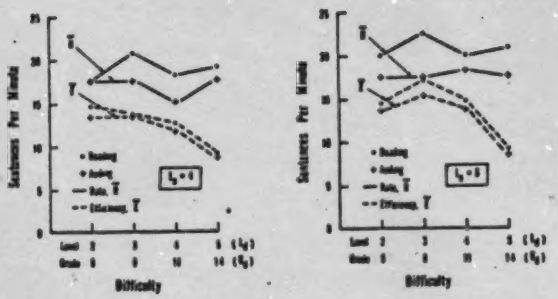


FIG. 7

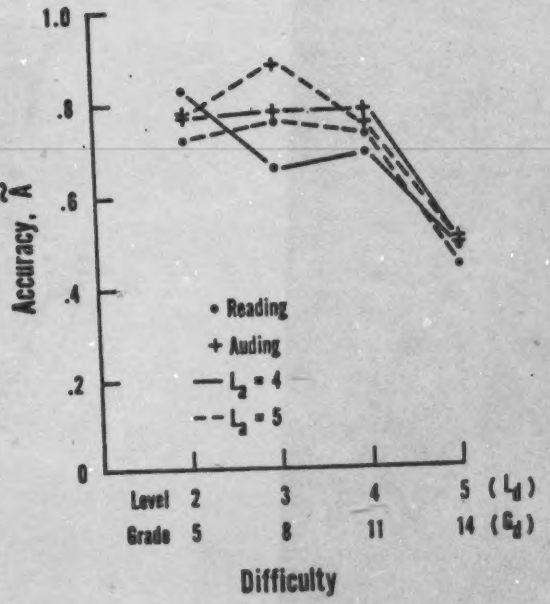

1.0
.8
.6
.4
.2
0
Accuracy, $\tilde{A}$
• Reading
+ Auding
— $L_2 = 4$
-- $L_2 = 5$
Level 2 3 4 5 (L_d)
Grade 5 8 11 14 (G_d)
Difficulty

FIG. 8

The Constancy of Reading Rate

Ronald P. Carver

University of Missouri-Kansas City

September 1977

The Constancy of Reading Rate [1]

Previous research has indicated that college students tend to read material of differing difficulty at the same rate (Carver, 1972; Carver, 1976a; Coke, 1974; Miller and Coleman, 1972). The constant rate is found when rate is controlled for word length by using letters per second or standard length words per minute rather than using actual words per minute. This suggests that individuals may be simply throwing their eyes down the line a constant number of character spaces regardless of the material, a view of reading advanced by Morton (1968). Rayner (1974) has referred to this type of eye movement as a "constant pattern" mechanism.

The constant pattern mechanism explains the constancy of reading rate across difficulty levels but it is in complete disagreement with the two factor theory of eye movements during reading advanced by Hochberg (1970). One of Hochberg's factors is termed peripheral search guidance (PSG) and the other factor is termed cognitive search guidance (CSG). PSG is described by Hochberg in the following way: "The low-acuity information picked up in the periphery of the eye can suggest to the optic search system where it must move its point of clearest vision in order to get a detailed view of some potentially interesting region [p. 75]." CSG is described in this way: "Knowledge about what he has seen so far should provide the observer with some hypotheses about where he should look in order to obtain further information [p. 75]." If individuals are able to use prior knowledge and thereby skip over redundant information, then it would seem that easy material, e.g., fourth grade difficulty material, should be read substantially faster than more difficult material, e.g., thirteenth grade difficulty material. The rival hypothesis that explains these data is that individuals simply throw their eyes a fixed distance, on the average, into the periphery; this is the constant pattern mechanism mentioned earlier.

The following research was designed to accomplish several purposes. First, it seemed important to determine whether the constant rate of reading across different difficulty levels could be replicated using different subjects, passages, and experimental conditions. Second, it also seemed important to determine whether or not individuals at reading levels other than the college level also read at a constant rate. Third, the rate of reading normal material was compared to the rate of reading material with extra spaces between words. If spaced material is read slower than normal material, this would provide evidence for a constant pattern mechanism being used during eye movements rather than the PSG and CSG mechanisms.

Method

Subjects. Students in grades 3 through 12 of a small town school system[2] provided 564 of the total of 666 individuals whose data were analyzed. The other 102 individuals were college students who were also participating in a much larger research project that will be described in detail elsewhere.

Materials. The 24 passages read by the subjects were taken from the 330, 100-word passages that were selected from curriculum materials and studied by Bormuth (1969). There were four passages selected randomly from each of six grade level difficulties -- 1, 4, 7, 10, 13 and 16 -- with the restriction that no passage be selected that was also being used as part of the larger experiment. The grade level difficulties of all 330 of the passages had been previously determined using the Rauding Scale of Prose Difficulty (Carver, 1974), as part of another study (Carver, 1975-76).

Design. There were two primary experimental prose conditions; one was Normal prose and the other was Spaced prose. The Normal prose was simply the

passage typed in a normal fashion using a standard IBM typewriter, 12 characters per inch, an extra half space between lines (1½ line spacing), an average six inch line width, with no hyphenated words at the margin. The Spaced prose was exactly the same as the Normal prose except an extra space was always added whenever there was normally a space, i.e., there were two spaces between words and three spaces between sentences.

All four of the passages at each of the six grade levels of difficulty were always on a single page, i.e., an 8½ x 14 sheet. The spaced version for the grade 16 material required all of the space on the page, i.e., 45 lines. The other grade levels and the normal spacing all required less than the full page because they were shorter in physical length. The shortest page was for the Normal, grade 1 material which required only 37 lines.

The Normal and Spaced versions of each of the 6 difficulty levels provided 12 conditions. For each of these 12 conditions there were two different orders of passage presentation providing a total of 24 different pages for reading. The two different order conditions were Forward and Backward, i.e., the Forward order of the four passages per page was randomly determined and then the Backward order was simply the reverse of the originally selected Forward order.

Each individual was presented a booklet containing 6 of the 24 possible pages. Each booklet started with a Normal page and then the Normal and Spaced pages alternated so that there were always three Normal pages and three Spaced pages per booklet.

Table 1 contains the overall design. Notice that the order of presentation of the six difficulty levels was counterbalanced using a Latin-square design for each of the two order conditions, i.e., Forward and Backward. The six numbered Forward conditions (1-6) and the six numbered Backward conditions (7-12) designated the booklet numbers that were affixed at the bottom of the booklet's cover

page. The booklets were stacked in sets of numbers 1-12 so that when they were distributed in each class of students, each of the 12 conditions would be distributed among the students equally.

Insert Table 1 about here

This design provided control over differences between passages within levels, order effects between Normal and Spaced presentation of material, and order effects for the presentation of the six difficulty levels.

Procedures. For the elementary and secondary students, the presentation of the booklet entitled the "Six Reading Rate Pages" was the only involvement the experimenter had with the students during the entire school year. For the college students, the presentation of this booklet was the first data that was collected at the outset of a full day of research participation.

The students were instructed at the outset that: (a) the purpose was to get an idea about their normal reading rate, (b) there would be no test on what they read, (c) the material should be read at the rate at which they normally read, (d) each of the six pages would contain material at a different difficulty level, (e) one minute would be allowed for each page, and (f) the word they were reading at the moment they were told to stop should be circled before they turned to the next page.

The elementary students were also told that much of the material would be very difficult for them to read and that if they found a word that they did not know that they should skip over it and continue reading. This instruction was given in an attempt to reduce possible anxiety effects.

Data Analysis. Almost all of the students who were administered the booklet had been given a standardized reading test, the National Reading Standards (Carver, 1977), one year previously. This test has been developed so as to

measure the reading ability of the individual along the same scale as the reading difficulty of the material, using the Rauding Scale of Prose Difficulty. A grade ability score of 5, for example, means that the individual is expected to be able to read and understand about 50% of the passages that are estimated to be at Grade 5 in difficulty.

Only the reading rate data of individuals who also had a grade ability score from the HRS were analyzed. Two forms of the HRS had been administered, so an average score was used to get a grade ability score for each individual. Then, one year was added to each score to provide an estimate of reading ability at the time the rate data was collected. Finally, the individuals were divided into groups containing three grades using previously prescribed conventions (see Carver, 1975-76 or Carver, 1977) for converting grade abilities, i.e., Level 1 ability was composed of individuals with grade abilities of 1-3; Level 2, Grades 4-6; Level 3, Grades 7-9; Level 4, Grades 10-12; Level 5, Grades 13-15; Level 6, Grades 16-18.

For the group of individuals who were in each of the six ability levels, the median reading rate was determined for both the Normal and Spaced versions at each grade difficulty. Since each individual received the passages at each of the six grade difficulties under only one of the Normal or Spaced conditions, individual differences between Normal and Spaced prose were not controlled at any specific grade difficulty. However, the individuals switched back and forth between Normal and Spaced for each succeedingly higher grade difficulty so individual differences were controlled over the entire range of grade difficulties.

For each individual reading a page, there were two different techniques used for determining reading rate. One method was simply to start at the first word on the page and count the actual number of words up to and including the word that had been circled on the page. A word was defined as what was contained between two blank spaces. Since one minute was provided for reading, this method

provided an estimate of rate in words per minute (wpm). The other method of estimating rate was to count the character spaces (including each blank space and each punctuation mark) up to and including the word that was circled, divide by six, and round off to the nearest word. This method, used in previous research (e.g., Carver, 1976), provided an estimate of rate in standard length words per minute (Wpm). Scoring keys were developed for determining both wpm and Wpm.

Results and Discussion

If the college level readers were simply throwing their eyes down a line an average number of character spaces with little or no regard for what was ahead, in a constant pattern, then they should read the Spaced material slower than the Normal material when rate is measured in actual words covered per minute, wpm. This is because the Spaced material is physically longer so it should take the reader longer to cover the same number of actual words. If college level readers actually read this constant character space way, the other measure of rate, in Wpm, should give exactly the same rate under both the Normal and Spaced conditions; this is because Wpm measures rate in character spaces. The results were just the opposite to what would be expected from the constant pattern mechanism.

Fig. 1 contains the reading rate of the college level readers, Level 5 (Grade 13, 14, and 15 abilities) who were reading materials at grades 1, 4, 7, 10, 13, and 16 difficulties. The rate is presented for the Normal and Spaced material under both the actual words per minute count, wpm, and the standard words per minute count, Wpm. It is quite obvious that the wpm measure of rate, represented by short dashed lines in Fig. 1, is almost exactly the same for both the Normal and Spaced materials. The rate of reading the Spaced version in Wpm, is consistently higher than the rate of reading the Normal version in Wpm--both of these rates in Wpm are represented by regular lines in Fig. 1. Thus, it is the wpm measure of rate which gives equivalent rates instead of Wpm as was predicted from the con-

stant pattern mechanism.

These data suggest that typical readers are able to anticipate the extra blank space between words and throw the eye just a bit further to correct for the extra distance. Thus, these data may be interpreted as evidence against the constant pattern mechanism and supporting the PSG mechanism advanced by Hochberg.

Insert Fig. 1 about here

The wpm results for Levels 2-5 are presented in Fig. 2. The Wpm data has not been included in Fig. 2 because there was not enough room in the figure; the Wpm data is not independent of the wpm data. (Note: Formulas have been presented by Carver, 1976a, for predicting Wpm from wpm given the difficulty level of the passage.) Notice that for reading ability Levels 3 and 4 the Normal and Spaced curves are almost perfectly coincident as they are for Level 5; the Level 5 data are exactly the same data as are presented in Fig. 1. The two Level 2 curves -- Grade 4-6 reading ability reading Normal and Spaced materials -- are almost perfectly coincident between Grade 1 and 4 difficulties; at Grade 7 difficulty and higher, where difficulty exceeds ability, the curves are more erratic.

Insert Fig. 2 about here

The Level 1 data were omitted from Fig. 2 because it was quite different and cluttered the figure. The Level 1 data is presented by itself in Fig. 3. Notice in Fig. 3 that there are notable differences between the two curves. For Grade 1, 3, and 7 difficulties, the spaced version was read about 30% faster than the normal version. It does not seem meaningful to try to interpret the inconsistent result at Grade 13 difficulty since this material is obviously far beyond the ability of these readers to comprehend.

Insert Fig. 3 about here

These results may be summarized as follows: For beginning readers, Level 1 readers, the extra space between words seems to have an effect, i.e., a positive or helpful effect. These readers may be less able to control their eye movements so that the extra space between words makes it easier for them to perceive the words while reading. For readers beyond Level 1, i.e., Grade 4 through 15 abilities, there seems to be an eye movement mechanism that allows these readers to adjust to the extra space between words rather automatically. This suggests that most readers do pick up word spacing in the periphery and use this information to direct where the eye will fixate next. These data may be interpreted as supporting the PSG mechanism advanced by Hochberg.

Rayner[3] has pointed out that a variant of the constant pattern mechanism could also explain these data. This mechanism would involve a gain control added to the constant pattern mechanism. For example, for the spaced material, the individual might simply lengthen, slightly, the constant amount the eye was being thrown down the line. The present data do not provide evidence against this variant of the constant pattern mechanism. However, Abrams and Zuber (1972) found that very few fixations were made upon large blank spaces between words in a passage, and Rayner (1974) reported that proportionately fewer fixations are made on the spaces between sentences. The more recent research of McConkie and Rayner (1975) also provides support for Hochberg's PSG mechanism. The present data seem to go further in reflecting upon the precision of the adjusting mechanism, whatever it is. These data suggest that individuals are able to adjust the length of the saccade so as to eliminate the effect of blank spaces between words which are as small as a single character space in width. Of course, eye movment data would be required to verify that it was in fact the saccade that was being lengthed to compensate for the extra space between words.

Since the readers were reacting to the two spaces between words no differently than the one space between words, the previously described Wpm method of measuring rate is not valid for the Spaced material.[4] Therefore, only the Wpm rate for the Normal material will be presented for Levels 1-5 -- see Fig. 4; the Level 5 data in Fig. 4 are exactly the same data as are presented in Fig. 1. An inspection of Fig. 4 reveals that when individuals at all levels of ability are reading material at a difficulty level that does not exceed their level of ability, then they tend to read the material at approximately the same rate. In Fig. 4, it may be noted that when the individuals at each ability level were reading material at a difficulty level that was higher than their ability level, the connecting lines have been designated by long dashes. Individuals at level 5 reading ability varied only between about 290 Wpm and 310 Wpm for materials at difficulty Levels 2, 3, 4, and 5. Material at difficulty Level 1 was read slightly faster at about 325 Wpm. These data from readers at Level 5 ability tend to replicate the college reader data collected by Miller and Coleman (1972), which was further analyzed by Carver (1976).

Insert Fig. 4 about here

General Discussion

There seems to be strong evidence that readers up through college level ability tend to read most material at the same rate as long as the length of the material is taken into account using a measurement technique such as standard length words per minute. These data also are in agreement with the recent data collected by Coke (1974) who concluded that both oral and silent reading rate ". . .remain constant across the entire range of difficulty of English prose when rate is measured in a unit smaller than a word [p. 409]." These data do not support the cognitive search guidance mechanism, CSG, of Hochberg. If there is an eye movement mechanism which is sensitive to context so that words carrying less meaning can be accurately predicted and skipped over, then there should be

a tendancy for the easier material to have been read faster than more difficult material. This was not the case. The idea that good readers typically hypothesize what words come next and then skip over certain words, supposedly the essence of Hochberg's CSG mechanism, also has been challenged by McConkie and Rayner (e.g., see Rayner, 1975). The data presented by these latter two researchers may be interpreted as providing additional evidence against the CSG mechanism of Hochberg's.

The present data do provide support for a recently developed theory, called rauding theory (Carver, 1977-78). Rauding theory is based upon the premise that most individuals typically read at a constant rate; the present data provide strong support for this fundamental aspect of the theory. Rauding theory holds that this constant rate is limited in one way by how fast the individuals can say the words to themselves and still comprehend the thought contained in a sentence. Thus, a better theoretical unit for studying the constancy of rate is probably syllables per min. However, since a unit of length involving letters is easier to measure and since it correlates so highly with a unit involving syllables, the standard length word is a more practical unit to use. (Coleman, 1971, found a correlation of .97 between average word length measured in letters and average word length measured in syllables.) Rauding theory contends that individuals internally articulate the words while they are reading normally and the constancy of this inner speech rate, in syllables per min., is one of the primary reasons why reading rate is constant.

Summary and Conclusions

These data replicated earlier findings which suggest that college students tend to read at a constant rate under normal or typical reading conditions. These data also indicate that the constancy of reading rate is not limited to students who read at the college level. Each level of reading ability appears to read different levels of material difficulty at approximately the same rate, at least as long as the difficulty level of the material does not exceed the ability level

of the reader.

Reading rate in actual words per minute was not affected by an extra space between words. Therefore, the constancy of reading rate is not easily explained by an eye movement mechanism which simply throws the eyes down a line a constant distance, on the average. The peripheral search mechanism, PSG, suggested by Hochberg can easily explain the constancy of reading rate for both the normal and spaced passages. These data then may be interpreted as supporting the theory that the low-acuity information picked up in the periphery of the eye is used by the visual system to suggest where it must move its point of clearest vision in order to get the clearest view of the words being read. These data also suggest that whatever mechanism is being used, it appears to be sensitive to visual stimuli as small as a single character space in width.

It would be difficult for the hypothesis testing approach to eye movements, as articulated in Hochberg's cognitive search mechanism, CSG, to explain the constancy of reading rate across differing difficulty levels under normal or typical reading conditions. There appears to be no data which requires the explanatory power of the CSG mechanism. What data there is, appears to be inconsistent with this theoretical mechanism, and this suggests that the mechanism should be closely scrutinized. These data, then, do <u>not</u> support the theory that under normal or typical reading conditions prior knowledge of the words and the information they carry suggest to the reader good hypotheses about where the point of clearest vision should be moved. Where the point of clearest vision is to be moved during normal reading seems to have a great deal to do with <u>where</u> the words are located but it does not seem to have much to do with <u>what</u> information the words carry.

With the data presently available, it seems reasonable to conclude that: (a) under typical reading conditions, reading rate is relatively constant and

does not vary with changes in the difficulty or redundancy of the material, (b) the hypothesis testing view of reading needs to be rejected, refined, or severly qualified, and (c) the premise of reading theory that rate is typically constant has considerable empirical support.

References

Abrams, S. G. and Zuber, B. L. Some temporal characteristics of information processing during reading. Reading Research Quarterly, 1972, 8, 40-51.

Bormuth, J. R. Development of readability analyses. U. S. Office of Education Final Report, Proj. No. 7-0052, Contract No. OEC-3-7-070052-0326, University of Chicago, March, 1969.

Carver, R. P. Evidence for the invalidity of the Miller-Coleman Readability Scale. Journal of Reading Behavior, 1972, 1, 42-47.

Carver, R. P. Manual for the Rauding Scale Qualification Test. Kansas City, Mo.: Revrac Publications, 1974.

Carver, R. P. Measuring prose difficulty using the Rauding Scale. Reading Research Quarterly, 1975-76, 9, 660-685.

Carver, R. P. Word Length, reading rate, and prose difficulty. Journal of Reading Behavior, 1976, 8, 193-203.

Carver, R. P. Technical Manual for the National Reading Standards. Kansas City, Mo.: Revrac Publications, 1977.

Carver, R. P. Toward a theory of reading comprehension and rauding. Reading Research Quarterly, 1977-78, 13, 8-63.

Carver, R. P. Optimal rate of reading. Unpublished manuscript, 1978

Coke, E. U. The effects of readability on oral and silent reading rates. Journal of Educational Psychology, 1974, 66, 406-409.

Coleman, E. B. Developing a technology of written instruction: Some determiners of the complexity of prose. In E. Z. Rothkopf and P. E. Johnson (eds.) Verbal learning research and the technology of written instruction. New York: Teachers College Press, Columbia University, 1971. [pp 155-204]

Gibson, J. J. Transcripted comment in J. F. Kavenagh (Ed.) Communicating by language: The reading process. Bethesda, Md.: National Institute of Child Health and Human Development, 1968.

Haber, R. N. Visual perception. In L. L. Elliot and J. A. Swets (Eds.) Psychology and the handicapped child. U. S. Department of Health Education and Welfare, DHEW. Publication No. (OE) 73-05000, 1974.

Hochberg, J. Components of literacy: Speculations and exploratory research. In H. Levin and J. P. Williams (Eds.), Basic Studies on Reading. New York: Basic Books, Inc., 1970.

Hochberg, J. and Brooks, V. Reading as an intentional behavior. In H. Singer and R. B. Ruddell (Eds.) Theoretical models and processes of reading. Newark, Del.: International Reading Association, 1970. [pp. 304-314]

McConkie, G. W., and Rayner, K. The span of the effective stimulus during a fixation in reading. Perception and Psychophysics, 1975, 17, 578-586.

Miller, G. R. and Coleman, E. B. The measurement of reading speed and the obligation to generalize to a population of reading materials. Journal of Reading Behavior, 1972, 3, 48-56.

Morton, J. Transcripted comment in J. F. Kavanagh (Ed.) Communicating by language: The reading process. Bethesa, Md.: National Institute of Child Health and Human Development, 1968.

Neisser, U. Cognitive psychology. New York: Appleton - Century - Crofts, 1967.

Rayner, K. The perceptual span and peripheral cues in reading. Unpublished doctoral dissertation, Cornell University, 1974.

Rayner, K. The perceptual span and peripheral cues in reading. Cognitive Psychology, 1975, 7, 65-81.

Smith, F. Understanding reading. New York: Holt, Rinehart, and Winston, 1971.

Footnotes

1 This research was supported in part by the National Institute of Education, Grant No. NE-G-00-3-0043. Fran Cohen and Debra Stein assisted with the analysis of the data.

2 The cooperation of the Pierce City, Missouri R-6 School District was appreciated--Mr. J. D. Smith, Superintendent; Mr. Earl Staponski, Elementary Principal; Mr. Donald Trotter, Secondary Principal.

3 Personal communication.

4 The method of using 6 character spaces per a standard length word is therefore only valid when the blank space between words is counted as one character space, no matter how large the space actually is.

Table 1 -

Design for the Presentation of the Passages
Under the Various Treatment and
Control Conditions

Booklet Number		Order of Presentation					
		1	2	3	4	5	6
Forward	Backward	Normal	Spaced	Normal	Spaced	Normal	Spaced
1	7	1	16	4	13	7	10
2	8	4	1	7	16	10	13
3	9	7	4	10	1	13	16
4	10	10	7	13	4	16	1
5	11	13	10	16	7	1	4
6	12	16	13	1	10	4	7

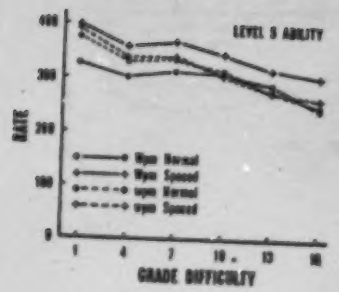
Figures

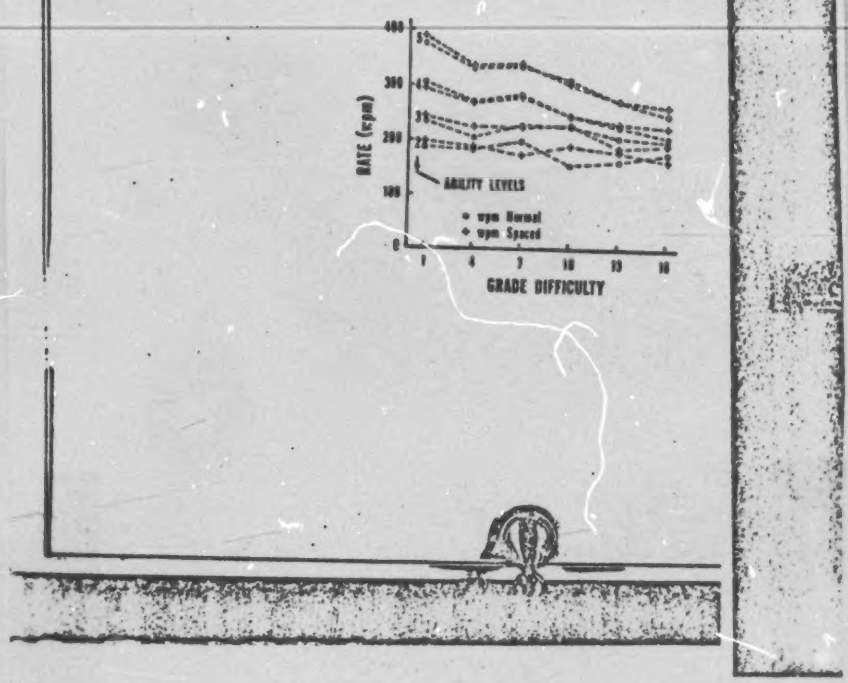
Fig. 1. Reading rate of college level readers, Level 5 ability, as a function of the difficulty of the reading material under two typographical conditions, Normal and Spaced, and using two different measurement techniques, standard length words per min., Wpm, and actual words per min., wpm.

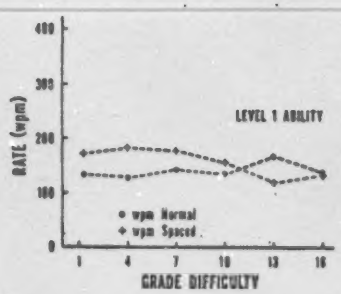
Fig. 2. Reading rate in actual words per min., wpm, as a function of the difficulty of the reading material for individuals at four different levels of reading ability who were reading both Normal and Spaced materials.

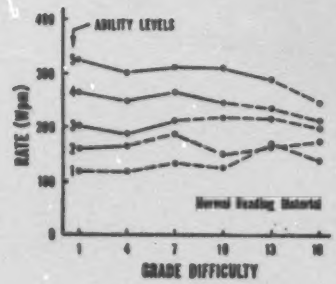
Fig. 3. Reading rate in actual words per min., wpm, as a function of the difficulty of the reading material for individuals at Level I ability who were reading both Normal and Spaced materials.

Fig. 4. Reading rate in standard length words per min., Wpm, as a function of the difficulty of the reading material for individuals at five different levels of reading ability who were reading the Normal reading material. (The long dashed lines indicate that the difficulty of the material exceeds the reading ability of the individual.)


LEVEL 9 ABILITY
RATE
GRADE DIFFICULTY


RATE (wpm)
400
300
200
100
0
ABILITY LEVELS
wpm Normal
wpm Spaced
1
4
7
10
13
16
GRADE DIFFICULTY


LEVEL 1 ABILITY
RATE (wpm)
400
300
200
100
0
wpm Normal
wpm Spaced
1
4
7
10
13
16
GRADE DIFFICULTY


ABILITY LEVELS
RATE (Wpm)
400
300
200
100
0
1
2
3
4
5
1
4
7
10
13
16
GRADE DIFFICULTY

Reading Comprehension under Normal and Speed Reading Conditions

Ronald P. Carver

University of Missouri-Kansas City

August 1977

Reading Comprehension Under Normal and Speed Reading Conditions[1]

Relatively precise predictions of the accuracy of reading comprehension at different rates are made in a recently presented theory of reading comprehension, called rauding theory (Carver, 1977-78). One important objective of the present research was to investigate the degree to which comprehension could be accurately predicted from rauding theory. In order to test the limits of the theory, comprehension at extremely high rates, i.e., speed reading rates, was compared to comprehension at low and normal rates. To determine whether the predictions are only accurate for certain ways of measuring comprehension, four very different measurement techniques were investigated to find out if they all reflect the same relationship between rate and accuracy of comprehension. Finally, since the predictions may only be accurate for certain reading purposes, the degree to which the reading purpose affected comprehension was also investigated.

Method

Introduction. The collection of these data came at the end of a more extensive data collection effort which is described in detail elsewhere (Carver, 1978).

Subjects. There were 20 groups of subjects with four to six subjects per group, i.e., a total of 102. The subjects were all college students who had responded to an advertisement in a university student newspaper. This advertisement stated that students would be paid $15.00 in cash, plus up to $13.00 in bonuses for their participation from 9:00 A.M. to 5:00 P.M. in an educational research project.

1

Passages. There were a total of 10 reading passages used. The passages were taken from the 330, 100 word passages that Bormuth (1969) sampled from curriculum materials in Grades 1-12 and college. As part of another study (Carver, 1975-76), the difficulty of the 330 passages had been graded using the Rauding Scale of Prose Difficulty (Carver, 1974b). From the 24 passages that had been rated as being Grade 10 in difficulty, 10 were randomly selected for use in the present study. The mean length of the 10 passages, in standard length words (W) was 101.8 words. A standard length word, W, was been defined as 6 letter-spaces (Carver, 1977-78.).

Rates. The passages were presented at 10 different rates--62.5, 125, 250, 500, 1000, 2000, 4000, 8000, 16,000, and 100,000 standard length words per min. (Wpm). Actually, the rate of reading was not directly manipulated; the time the passage was presented was directly manipulated. The amount of time that each passage was presented at each rate was determined in three steps. First, the number of standard length words, W, in each passage was counted. Next, the number of standard length words, W, was divided by the rate desired in Wpm to give the length of time, in min., that the passage would be presented. Finally, the length of time, in seconds, was multiplied by 18 to give the number of motion picture frames that would be shot of the passage. Since the motion picture projector had been modified to run at a rate of 18 frames per sec., this procedure had the effect of presenting each passage for the specific length of time required to provide the average rate desired. Under the 100,000 Wpm condition, the passage was presented in only one frame.

Design. There were 10 different films made. Each film contained the 10 passages at the 10 different rates. The passages, rates, and presentation orders were counterbalanced over the 10 films. The first 10 groups received the 10 films under one purpose condition and the second 10 groups received the same 10 films under another purpose condition.

Standardized Test. As part of the larger data collection effort, the National Reading Standards (NRS) had been administered earlier to these 102 subjects. This test is a standardized, but reading test which provides a reading efficiency score (Carver, 1977). This score may be converted into a Grade Ability score and a Reading Rate score. The median Grade Ability score for these 102 subjects was 13.5 and the corresponding median Reading Rate score was 305 Wpm. The 13.5 Grade Ability may be interpreted as indicating that the individual at the average of this group could be expected to be able to read and understand about 50 percent of all the reading materials that are rated at 13.5 Grade Difficulty, according to the Reading Scale of Prose Difficulty mentioned earlier. The 305 Wpm Reading Rate may be interpreted as indicating that the individual at the average of this group would be expected to have a normal reading rate of 305 Wpm, i.e., this individual would be expected to comprehend sentences in passages most efficiently at a rate of 305 Wpm (see Carver, 1976).

Experimental Measures. There were four experimental measures--Understanding Judgments, Missing Verbs, Best Titles Test, and Sentence-Halves test.

Immediately after having been presented a passage, the subjects were asked to estimate the percentage of the total number of complete thoughts in a passage that they understood, i.e., the Understanding Judgments measure. The subjects had had a great deal of instruction and practice in making these judgments as part of the earlier data collection efforts. This estimation procedure is described in detail in a number of earlier publications (Carver, 1973; Carver, 1974a; Carver, 1975). This measure involved placing a circle around one of the following numbers to reflect percentage of understanding--0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100. This measure was used to provide an estimate of the accuracy with which the sentences in each passage had been comprehended, i.e., a percent comprehension estimate.

As soon as the Understanding Judgment had been made, the subject was supposed to circle either Yes or No indicating whether or not he or she thought a verb had been deleted. In five of the ten passages, a verb had been deleted from one of the sentences in the passage. The choice of which sentence to delete a verb was random except there was a restriction that the five sentences selected must represent each fifth of a passage. This restriction prohibited the missing verbs from falling into any one particular section of a passage. This Missing Verbs measure was included so as to have a completely objective indication of the accuracy of comprehension of the sentences, i.e., percent comprehension.

After all 10 of the passages had been presented, the two tests on the passages were administered. The Sentence-Halves Test was always administered before the Best Titles Test.

The Sentence-Halves Test contained 40 halves of sentences. Two sentence-halves had been randomly selected from each of the 10 passages, one first half and one second half from each passage. These 20 sentences-halves were randomly mixed with 20 other sentence-halves that had been randomly selected from 10 other passages. Each one of the 10 other passages was matched on subject matter and difficulty as close as possible to its experimental counterpart. The task for the subject was to circle Yes or No in front of each sentence-half designating whether or not the subject thought he or she had seen the sentence-half before in the 10 passages which had been presented. The subjects were given eight min. to work on this test and this was ample time for more than 90 percent of the subjects to finish. This measure was a recognition type of measure not necessarily connected with the comprehension of the sentences in a passage. It would seem to be possible for an individual to quickly skim over a passage and remember phrases and words without comprehending the sentences; this measure was designed to reflect that type of ability.

The Best Titles Test consisted of 10 multiple choice questions. Each question consisted of five possible titles for each of the 10 passages. The task for the individual was to choose the single title out of the five available which was best for the passage. The subjects were given four min. to work on the 10 items on this test, and this time limit seemed to be more than ample for over 90 percent of the subjects. The purpose of this measure was to reflect the extent to which the subjects could quickly skim over a passage and get the gist of the passage.

On both the Sentence-Halves Test and the Best Titles Test, the order of the items was the same for all groups. However, since the order of presentation of the passages and rates was different for each of the 10 films, this had the effect of balancing out order effects on the test.

Procedures. At approximately 4:25 P.M., after an 8 min. rest break, the present experiment was begun. The subjects were told that they would be presented 10 passages for reading and the rates would vary from extremely slow 62.5 words per min., to extremely fast, 100,000 words per min. They were also told that the rate at which each passage would be presented would appear on the screen immediately before they were presented the passage to read. This rate appeared on the screen at the outset for 1 sec. and then there was a 1 sec. pause prior to the presentation of the passage. The four different experimental measures were explained to them and then they were given a passage at 250 Wpm for practice. After the practice passage had been presented, they: (a) marked their percent of understanding, (b) marked yes or no indicating whether they thought a verb was missing, (c) answered four Sentence-Half items, and (d) answered a Best Titles item. It was carefully explained to them that the percent understanding and verb missing tasks would be presented immediately after each passage, as it was with the practice trial, while the Sentence-Half Test and the Best Titles Test would be given at the end, unlike the practice trial.

Groups 1-10 were told that the most important thing they should be doing while they were reading was to try to understand enough of the passage to answer correctly the questions on the Best Titles Test. They were told that they would be given: (a) 10¢ for each correct answer on the Best Titles Test, (b) 3¢ for each correct answer to the Missing Verb questions, (c) 2¢ for each correct answer on the Sentence-Halves Test.

Groups 11-20 were told that the most important thing they should be doing while they were reading was to try to determine whether or not there was a verb missing in the passage. They were told that they would be given: (a) 10¢ for each correct answer on the Missing Verb question, (b) 3¢ for each correct answer on the Best Titles Test, and (c) 2¢ for each correct answer on the Sentence-Halves Test.

Equipment. Each of the 10 reading films were projected onto a standard 5' x 6' motion picture screen situated about 8' from a 16 mm projector that had been modified to run at a rate of 18 frames per second. Since there were only 6 subjects, at most, each subject was able to move to a viewing position that was comfortable.

Data Analysis. For each measure and for each of Groups 1-10, the mean of the scores from the 4 to 6 subjects in each group was calculated at each rate. Then, the mean of the 10 means at each rate was calculated. This analysis was repeated for Groups 11-20. The three objective measures were corrected for guessing, i.e., Rights minus Wrongs for Missing Verbs and Sentence Halves, and Rights minus $\frac{1}{2}$ Wrongs for Best Titles.

Results and Discussion

Fig. 1 contains four graphs relating average reading rate to each of the four measures of the accuracy of comprehension. Each graph contains two curves.

One curve is for Groups 1-10 which had the primary goal of determining the best title for each passage, a very general purpose. The other curve is for Groups 11-20 which had the primary goal of determining whether or not a verb was missing from each passage, a very specific purpose. Notice that the scale is different for different rates, i.e., for rates 62.5 to 1000 Wpm one unit equals 100 Wpm, for rates 2,000 to 8,000 Wpm one unit equals 1,000 Wpm, and for rates 16,000 to 100,000 Wpm one unit equals 10,000 Wpm. This change in the scale helps to accent the differences at the lower rates.

Insert Fig. 1 about here

For the Understanding Judgments measure, there appears to be no substantive difference between the two groups. For the Missing Verbs measure, there does appear to be a consistent and substantial difference (e.g., 22 percentage points at 250 Wpm) in favor of the groups which were instructed to focus upon the missing verbs. For the Best Titles Test measure, there appears to be no consistent, substantive difference between the two curves. For the Sentence-Halves measure, there appears to be consistent differences in favor of Groups 1-10 which focused on determining the Best Titles (15 percentage points at 250 Wpm).

The groups which focused upon the general purpose of being able to discern the best titles for the passages did not seem to do any better on the Best Titles Test than the groups which focused upon the specific purpose of detecting whether or not a verb was missing. This suggests the lack of an ability to get the gist of a passage by skimming, or speed reading, the passage.

The groups which focused upon detecting the missing verbs could do a better job on this task, but it did not seem to put them at a disadvantage on the more general task of discerning the best titles for the passages. This suggests that

92A

individuals can focus upon comprehending each sentence without losing the ability to get the gist of a passage.

The general purpose seemed to have resulted in a somewhat better recognition of which sentence-halves had been presented in the 10 passages. Given the preceding results and interpretations, this result is difficult to explain. There seems to be no good reason why the general purpose group should have been able to remember which sentence-halves had been seen before if they had no better comprehension of the individual sentences and had no better acquisition of the gist of the passage.

Probably the most succinct way to interpret these results is: the purposes in this experiment seemed to have little effect upon the comprehension of the sentences. Although there seemed to be some effects associated with certain purposes, they do not appear to be large enough to justify the consideration of purpose as a major variable in this experiment. From certain hypothesis testing theories of reading (e.g., see Hochberg, 1970, or Haber, 1974), and from certain advocates of speed reading such as Evelyn Wood Reading Dynamics (see Carver, 1972) it could be inferred that some individuals are able to rapidly read through a passage and get the general gist of the passage when general gist is required. These data do not support this position. These college students, who, as a group should be among the best readers in the world, were not able to focus upon getting the gist of a passage and do any better on a measure of gist than students who were focusing upon the individual words to see if they formed a complete thought.

The data in Fig. 1 may also be interpreted in relationship to the question of just how much was comprehended at the high rates. Notice that the Understanding Judgments were around 80-100% at the low rates, 62.5 to 250 Wpm. However, at 500 Wpm these understanding ratings dropped precipitously to about 40% and at 1000 Wpm they were around 15-20%. Beyond 1000 Wpm these ratings rapidly approached zero. This suggests that there is little comprehension at the high rates. Furthermore, it should be remembered that rate was not actually manipulated at all; the time allowed to read was manipulated. If the rate had actually been manipulated, there is no doubt but that percent comprehention would have been even lower (see Carver, 1978). The data from the Missing Verbs test further supports the preceding interpretation. The accuracy of detection of the missing verbs was around 60-100% at the low rates, 62.5 to 250 Wpm. However, at 500 Wpm the accuracy dropped precipitously to only 20-30%. At 1000 Wpm, the accuracy was 10%, and beyond 1000 Wpm the accuracy scores seem to fluctuate around the chance score of 0%.

It is not proper to interpret the scores on the Best Titles Test and the Sentence-Halves Test on the same absolute percentage basis as were the Understanding Judgments and Missing Verbs measures above. The percent correct on the Best Titles Test depended upon the subjective judgments of the test developer, i.e., it would have been possible to develop more obviously incorrect alternatives and thus raise the percent correct. Furthermore, both of these tests were administered after a certain undeterminable amount of forgetting had occurred, i.e., both were administered after all 10 passages had been presented. However, it will be demonstrated later that these latter two measures reflect the same low degree of comprehension at high rates when they are placed upon the same scale as the two former measures.

In order to determine the degree to which the four different measures reflect the same relationship between rate and accuracy, the four measures were intercorrelated. Since there seemed to be little difference between Groups 1-10 and Groups 11-20, their two means at each rate were combined, i.e., their two curves were collapsed for each measure by calculating the mean of the pair of means involved. These resulting overall mean scores at each of the 10 rates were intercorrelated for the four variables. The resulting six correlations ranged from .92 to .99, i.e., very high relationships. The standard error of estimates in predicting the Understanding Judgments using the appropriate regression equations were very small, i.e., 5, 11, and 3 percentage points for the Best Titles Test, Missing Verbs, and Sentence-Halves Test, respectively. However, it was noted from a scatter plot of these relationships that the low scores on the measures, except for the Understanding Judgments measure, seemed to fluctuate radically. This was interpreted, as mentioned before, as simply fluctuations around chance at the rates greater than 1000 Wpm. Therefore, the same correlations were calculated again, but this time only the lowest five rates were used--62.5 to 1000 Wpm. These results are presented in Table 1. Notice that the six correlations now range only from .95 to 1.00. Again, the data for each objective measure were used to calculate regression equations for predicting Understanding Judgments, and the corresponding standard error of estimates were 2, 4, and 1 percentage points for the Best Titles Test, Missing Verbs, and Sentence-Halves Test, respectively. These are extremely low errors of prediction. These data may be interpreted as indicating that all four measures reflect the same relationship between rate and accuracy of comprehension.

Insert Table 1 about here

The final question to be investigated concerns the degree to which comprehension may be predicted using rauding theory. Rauding theory purports to be able to predict the accuracy of reading comprehension as long as values of the following variables can be accurately estimated: the rauding rate of the individual, R_r; the rauding accuracy of the individual for the passage, A_r; the length of the passage, T_p; and the time spent reading, t. A table of values is presented with the theory (see Carver, 1977-78) that can be used to estimate A_r from estimates of ability, G_a, and difficulty, G_d. This same table can be used to estimate R_r from an estimate of ability, G_a. In the present research, the grade level of ability of the group was estimated earlier as 13.5 so G_a is estimated to be 13.5. The difficulty level of the passages was given earlier as 10 so G_d is estimated to be 10. Entering the above mentioned table with these estimates of G_a and G_d, provides a .78 estimate for A_r and an 18.3 standard length sentences per min. (Spm) estimate for R_r. (Note: As prescribed in the theory, Spm may be converted into Wpm by multiplying by 16.67, i.e., 18.3 Spm = 305 Wpm). The mean length of the 10 passages, as noted earlier, was 101.8 standard length words, and this may be converted into an estimate of the number of thoughts, T_p, by dividing by 16.67, as prescribed in the theory. Therefore, T_p is estimated to be 6.11 thoughts. The time, t, spent reading each passage was actually the variable that was manipulated in the experiment; it was originally determined by dividing the number of standard length words, W, by the desired Wpm for each passage. The mean time, t, for all 10 passages at each rate was determined by dividing the average passage length, 101.8 Words, by each rate in Wpm. For example, the average length of time the passages were presented at the 250 Wpm rate was determined by dividing 101.8 by 250 to give .41 min. or 25 sec. Thus, all the necessary values for predicting accuracy of comprehension (A) may be accurately estimated in this experiment.

According to rauding theory, there is one equation for predicting the accuracy of comprehension, A, when the individual has been given less than enough time to

finish reading the passage once, i.e.,

$$A = A_r R_r \frac{t}{T_p}, \text{ when } t \leq \frac{T_p}{R_r}, \tag{1}$$

and another equation for predicting the accuracy of comprehension, A, when the individual has been given more than enough time to finish reading the passage once, i.e.,

$$A = \frac{1}{1 + \frac{1}{R_r}\left(\frac{1}{A_r} - 1\right)\frac{T_p}{t}}, \text{ when } t \geq \frac{T_p}{R_r}. \tag{2}$$

Since

$$\frac{T_p}{R_r} = \frac{6.11}{18.3} = .33, \tag{3}$$

this means that the equation for predicting the accuracy of comprehension, A when the time allowed for reading the passage, t, is less than or equal to .33 min. is

$$A = (.78)(18.3)\frac{t}{6.11} = 2.34t, \tag{4}$$

and when the time allowed for reading, t, is equal to or greater than .33 min., the prediction equation is

$$A = \frac{1}{1 + \frac{1}{18.3}\left(\frac{1}{.78} - 1\right)\frac{6.11}{t}} = \frac{1}{1 + \frac{.094}{t}}. \tag{5}$$

In rauding theory, the accuracy of comprehension of a passage is a proportion--the number of thoughts comprehended or understood divided by the total number of thoughts in the passage. As noted earlier, Understanding Judgments provide a relatively accurate estimate of this variable when the percent values are divided by 100 to convert them into proportions. The other three measures may be transformed into similar accuracy of comprehension estimates by use of the three regression equations which allow the Understanding Judgment measure to be predicted from each of the three measures. These three regression equations

were calculated from the means, standard deviations and correlation coefficients presented in Table 1. Thus, the three measures besides the Understanding Judgments may be rescaled, using a linear transformation which does not affect the shape of the curves, so as to put all four measures on comparable accuracy of comprehension scales.

Fig. 2 presents accuracy of comprehension as a function of the average time allowed for reading each passage in minutes. The theoretical prediction curve presented in Fig. 2 was derived by substituting into Equations 4 and 5, values of t varying from 0 to 1.6 min. Fig. 2 also contains the accuracy of comprehension for the four measures, using the actual proportions for the Understanding Judgments and using the three regression equations, described above, to predict accuracy of comprehension for the other three measures. Only the data for the lowest five rates, 62.5 Wpm (1.63 min.) to 1000 Wpm (0.10 min.) are plotted in Fig. 2. However, the other 5 scores for each of the three measures were also predicted using the same regression equations. The mean of these 15 accuracy estimates was .01; this value compares very favorably with the mean of the five corresponding Understanding Judgment proportions, i.e., .02. Thus, this estimation procedure provides internally consistent results. It may be inferred from the .01 and .02 estimates that the accuracy of comprehension was near zero at average rates over 1000 Wpm.

Insert Fig. 2 about here

The most phenomenal result, however, was the extremely close correspondence between the accuracy of comprehension predicted from rauding theory, i.e., the theoretical prediction curve, and the actual estimates of accuracy which came from the three objective tests and the Understanding Judgments. For each of the four measures, the five absolute differences between the actual value and

the predicted value were calculated. The mean of these five absolute differences was extremely low for each of the four measures--Understanding Judgments, .03; Missing Verbs, .04; Best Titles Test, .05, and Sentence-Halves Test, .03. Thus, it can be seen that reading theory predicts accuracy of comprehension in an extremely accurate fashion, no matter which of the four measures are used. These data also may be summarized as follows: The theoretical estimates of comprehension accuracy were within about 3.5 percentage points of the empirical estimates of comprehension accuracy.

Before terminating the discussion of these results, it should be emphasized that the three objective measures provided almost exactly the same results as the subjective judgments of the subjects themselves. In the previously noted research where similar subjective judgments have been used, the judgments immediately preceded an objective test on the passage. Thus, it might have been suspected that the high correlation between the objective measures and the subjective measures in that research was artifactually inflated. For example, the subjects might have estimated how well they thought they would do on the test and then let this judgment influence their understanding rating. This hypothesis does not seem credible in light of the present research, however. Consider the almost perfect prediction of the Understanding Judgment means from the Missing Verbs means. Although the Understanding Judgment measure varied continuously from 0-100 percent, the Missing Verbs measures was dichotomous, Yes or No. This high predictive accuracy could only result from the calculation of the Missing Verb means of a large group of subjects, i.e., so that the predicted percent comprehension values were not dichotomous for the group but varied continuously from 0-100. Therefore, it was impossible for the high relationship between the subjective measure and the objective measure to have resulted from the kind of intra-subject contamination factor mentioned above.

Summary, Conclusions, and Implications

These data provide strong support for rauding theory. It appears that these individuals simply read the passages at their normal rate and in their normal fashion until the end of the time limit. It does not appear that different purposes during reading, at least given the range of purposes used in this research, had much effect upon what the individual did while reading. These data provide no support for those who would contend that purpose for reading has a large effect upon comprehension in relatively normal or typical reading situations. Those individuals who were primarily trying to get the gist of the passage were no better able to accomplish this end than those individuals who were primarily reading each sentence to determine if a verb was missing. There were some differences attributable to the particular purpose but the commonality of the results were more outstanding than the differences.

When the research was designed, it was anticipated that there would be some differences among the four measures used. However, one of the most significant findings of this research was how the four measures provided almost perfectly replicated results. This suggests that rauding, i.e., comprehending sentences in passages (see Carver, 1977-78), is probably a prerequisite type of factor in the reading process. Probably, the ability to detect missing verbs, the ability to get the gist, and the ability to recognize material that has been seen before, are all side effects of the more basic effect of reading, i.e., comprehending the thoughts represented by the sentences in the material.

These data seem to provide strong support for that part of rauding theory which contends that reading comprehension depends primarily upon two factors, the rauding rate of the individual, R_r, and the rauding accuracy of the individual for the passage, A_r. When the individual is given more time to read a passage, the individual will

comprehend more, i.e., accuracy of comprehension will increase. Given the extremely low degrees of comprehension associated with the high rates (500 Wpm and greater) no matter what measure is used, it is amazing that certain researchers (e.g., see Neisser, 1966) and certain rapid reading advocates (e.g., see Smith, 1971) seem to maintain that 1000 Wpm reading with high comprehension is an uncontested fact. It appears more reasonable to contend that even mature college readers have an optimal rauding rate, R_r, around 250 to 350 Wpm, and a certain rauding accuracy, A_r, for a certain passage of prose material. It appears that it is unreasonable to contend that simply forcing mature readers to read faster, i.e., the usual strategy of speed reading courses, will thereby increase R_r. Forcing individuals to cover materials at rapid rates may be justifiable in an effort to train better skimming practices. However, it does not seem reasonable to expect such training to increase R_r. The second primary way to increase the accuracy of comprehension per a fixed amount of time spent reading is to decrease the level of difficulty of the material. Thus, speed reading courses which include a great deal of practice upon extremely easy materials (e.g., see Carver, 1971) are capitalizing upon this factor. Also, the use of a preview or overview of the material in speed reading courses is one way that time, t, can be increased and thereby increase accuracy, A.

It would seem to be incumbent upon the individuals who operate commercial speed reading courses and college speed reading courses to demonstrate that the rate at which sentences in passages can be read and understood, R_r, can be increased by such training. If R_r can be increased by such training, then large theoretical and applied resources ought to be expended upon this area immediately. If mature readers, such as college students, really can be trained to double or triple the rate at which they can transform the words in sentences into the complete thoughts which they represent, then this training should be made universal. On the other hand, if it is impossible for most mature readers to learn to read

and understand sentences in passages at rates greater than about 400 Wpm, as suggested in rauding theory, then the individuals who conduct speed reading courses should scrutinize their advertisements more closely. Certainly, those individuals who conduct reading courses in association with colleges and universities should attempt to collect solid experimental data which supports their claim that individuals can be taught to increase the rate at which they comprehend sentences in passages during reading. Otherwise, they might consider advertising that they are teaching skimming skills and studying strategies without suggesting to naive students that they can be trained to read faster.

References

Bormuth, J. R. Development of readability analyses. U.S. Office of Education Final Report, Project No. 7-0052, Contract No. OEC-3-7-070052-0326, University of Chicago, March, 1969.

Carver, R. P. Sense and nonsense in speed reading. Silver Spring, Md.: Revrac Publications, 1971.

Carver, R. P. Speed readers don't read; they skim. Psychology Today, 1972, August, 22-30.

Carver, R. P. Understanding information processing, and learning from prose materials. Journal of Educational Psychology, 1973, 64, 76-84.

Carver, R. P. Measuring the primary effect of reading: Reading-storage technique, understanding judgments, and cloze. Journal of Reading Behavior, 1974, 6, 249-274. (a)

Carver, R. P. Manual for the Rauding Scale Qualification Test. Kansas City, Mo.: Revrac Publications, 1974. (b)

Carver, R. P. Comparing the reading-storage test to the paraphrase test as measures of the primary effect of reading. Journal of Educational Psychology, 1975, 67, 274-284.

Carver, R. P. Measuring prose difficulty using the Rauding Scale. Reading Research Quarterly, 1975-76, 11, 660-685.

Carver, R. P. Technical Manual for the National Reading Standards. Kansas City, Mo.: Revrac Publications, 1977.

Carver, R. P. Toward a theory of reading comprehension and rauding. Reading Research Quarterly, 1977-78, 13, 8-63.

Carver, R. P. Optimal Rate of Reading. Unpublished manuscript, 1978.

Weber, Ralph Norman Visual perception. In Psychology and the handicapped child. U. S. Office of Education, DHEW Pub. No. (OE) 73-05000, U. S. Government Printing Office Stock No. 1780-01219, 1974. [Pp. 40-72]

Hochberg, J. Components of literacy: Speculations and exploratory research. H. Levin and J. P. Williams (Eds.) Basic studies on reading. New York: Basic Books, 1970.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1966.

Smith, F. Understanding reading. New York: Holt, Rinehart and Winston, 1971.

Footnote

1 This research was funded by the National Institute of Education, Grant No. NE-G-00-3-0043. The author wishes to thank Fran Cohen and Debra Stein for their help in the data collection and data analysis.

Table 1

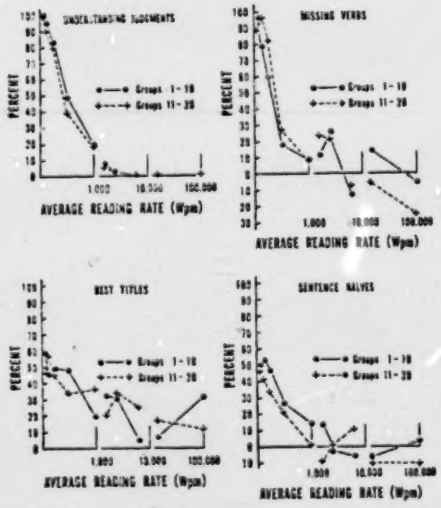
Intercorrelations, Means, and Standard Deviations for the Four Measures of Comprehension.

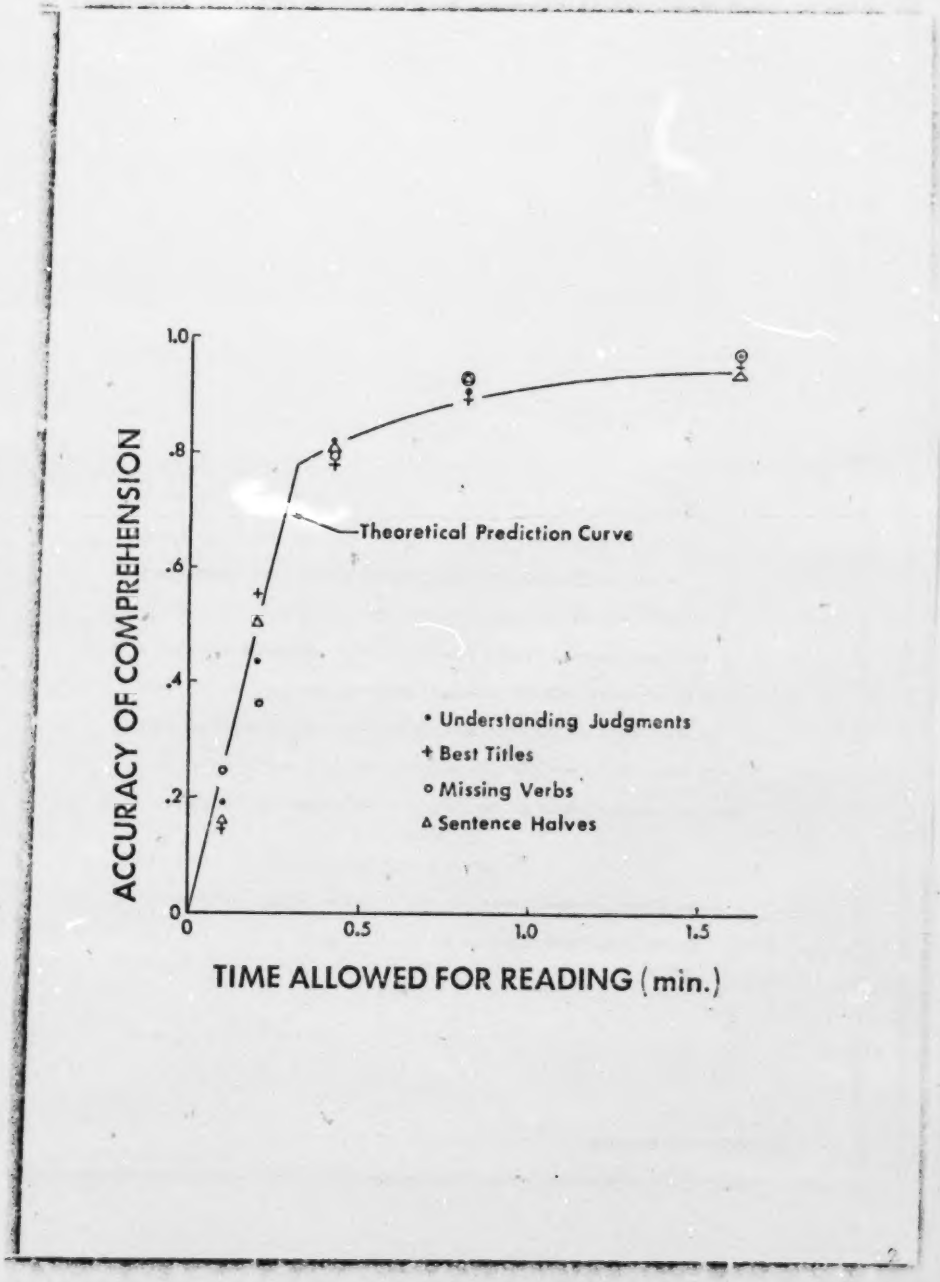
		1	2	3	4	Mean	S.D.
Understanding Judgments	1	1.00	.93	.99	1.00	.67	.31
Best Titles	2		1.00	.95	.99	.44	.09
Missing Verbs	3			1.00	.98	.56	.34
Sentence Halves	4				1.00	.34	.15

Figures

Fig. 1. Four measures of comprehension, in percent, as a function of the average reading rate for two different sets of groups which had different primary purposes for reading.

Fig. 2. Accuracy of comprehension as a function of time allowed for reading, in min., for all four measures of comprehension. Each objective measure reflects a predicted accuracy of comprehension score using the understanding judgments as the criterion in a regression equation. Also presented is the theoretically predicted relationship between accuracy and rate for this group of individuals who are estimated to have a reading ability of Grade 13.5, and who are reading material estimated to be at Grade 10 difficulty.


PERCENT
Groups 1 - 10
Groups 11 - 20
AVERAGE READING RATE (Wpm)
MISSING VERBS
TEST TITLES
SENTENCE HALVES


ACCURACY OF COMPREHENSION
1.0
.8
.6
.4
.2
0
0
0.5
1.0
1.5
TIME ALLOWED FOR READING (min.)
Theoretical Prediction Curve
• Understanding Judgments
+ Best Titles
o Missing Verbs
Δ Sentence Halves

MEASURING ABSOLUTE AMOUNTS OF READING COMPREHENSION USING THE RAUDING RESCALING PROCEDURE

Ronald P. Carver

University of Missouri-Kansas City

MEASURING ABSOLUTE AMOUNTS OF READING COMPREHENSION USING THE RAUDING RESCALING PROCEDURE[1]

Measuring absolute amounts of reading comprehension is not a simple matter. Easy comprehension questions can be written on material that is difficult to comprehend, and conversely, difficult questions can be written on material that is easy to comprehend. Answering 75 percent of the test questions correctly does not necessarily mean that 75 percent comprehension has been achieved. It may mean that many of the test questions were very easy, e.g., questions that could be answered without ever reading the passage. Or, it may mean that some of the questions were extremely difficult so that no one could get a 100 percent score even though they actually comprehended 100 percent of the passage.

Before discussing this measurement problem further, an attempt will be made to define the concept of reading comprehension with enough precision to justify further work on quantification. A theory has recently been developed (Carver, 1977-78) which advocated that the type of comprehension which occurs during reading involves three primary component variables--accuracy (A), rate (R), and efficiency (E). Accuracy, A, is defined as the number of complete thoughts, or sentences, that have been comprehended, or understood, per the number that are presented in the reading passage. Rate, R, is defined as the number of complete thoughts, or sentences, that were presented per the amount of time spent reading. Efficiency, E, is the number of thoughts, or sentences, that have been comprehended per the amount of time spent reading. Efficiency is also, by definition, the product of accuracy and rate,

$$E = AR. \qquad (1)$$

The focus of the present article will be upon measuring absolute amounts of the accuracy of comprehension, A. Notice that A is a proportion which is .00 when there is 0% comprehension--when none of the thoughts in a passage have been understood or comprehended. Accuracy, A, is 1.00 when there is 100% comprehension--when all of the thoughts in a passage have been understood or comprehended. Therefore, when an empirical indicant of the degree of comprehension is constructed, i.e., a test or a measure of comprehension, the problem is how to rescale the observed indicant of accuracy ($\tilde{A}$) so that it reflects absolute amounts of A.[2] Although $\tilde{A}$ is actually a proportion, it will often be referred to as "percent comprehension" because this is generally the terminology used to refer to an empirical measure of comprehension. For example, suppose a measure or an indicant of the accuracy of comprehension was administered to a group of individuals who had never read a passage and they got 40 percent of the questions correct. The percent comprehension, $\tilde{A}$, would be 40 percent, or .40 expressed as a proportion. A rescaling procedure would need to be developed so that a .40 value for $\tilde{A}$ was rescaled to provide a .00 estimate of A.

The primary purpose of this article will be to describe a procedure for rescaling percent comprehension measures, $\tilde{A}$, so that they reflect estimates of absolute degrees of comprehension, A. After the rescaling procedure has been described, it will be illustrated using empirical data. Under conditions where it is reasonable to assume that all comprehension measures should produce exactly the same result, it will be shown that quite different techniques for measuring reading comprehension produce quite different values of $\tilde{A}$ at each rate, but when they have been objectively rescaled to reflect A the different techniques all produce approximately the same estimated values of A at each rate.

Theoretical Rationale for the Rescaling Procedure

Rauding Theory

In the newly developed theory of reading comprehension, called rauding theory, the relationship between A and the inverse of rate, 1/R, is linear up to the point where the individual has had time to finish reading the passage once. The inverse of rate, 1/R, is the amount of time in min. allowed to read each complete thought or sentence. Therefore, in rauding theory the number of thoughts, or sentences, comprehended is considered to increase linearly with increases in the amount of time allowed for reading. The individual is assumed to read the passage at a constant rate called the rauding rate (R_r), and therefore there is a certain amount of time required to read the passage at the rauding rate (t_r). The accuracy of comprehension, A, at time t_r is defined as the rauding accuracy (A_r). The above theoretical relationships have been described mathematically as

$$A = (A_r R_r)\frac{1}{R} \text{ when } t \leq t_r. \tag{2}$$

When the time allowed to read a passage is greater than that required to read it once, $t > t_r$, then the interrelationships among accuracy, A, and rate, R, become more complex. A linear relationship between 1/E and 1/R is postulated with a slope of unity (1.00) and an intercept (I) that varies inversely with A_r and R_r. That relationship is as follows:

$$\frac{1}{E} = (1.00)\frac{1}{R} + I, \text{ and} \tag{3}$$

$$\frac{1}{AR} = (1.00)\frac{1}{R} + I, \text{ so that} \tag{4}$$

$$A = \frac{1}{1.00 + IR} \text{ when } t \geq t_r. \tag{5}$$

When $R=R_r$ and $A=A_r$, the solution for I from Equation 5 is

$$I = \frac{1}{R_r}\left(\frac{1}{A_r} - 1\right). \tag{6}$$

According to the theory, the value of R_r for an individual can be estimated from a standardized test score. The value of A_r can be estimated from the joint knowledge of the difficulty level of the passage (as estimated from a standardized procedure) and the ability level of the individual (as estimated from a standardized test). Given these estimated values of A_r and R_r, Equation 6 can be used to estimate i. Then, all values are available for estimating the accuracy of comprehension of a passage, A, using either Equation 2 or Equation 5, provided that the average rate, R, the passage is presented is also known.

The Theoretical Rescaling Procedure

The essential assumption underlying the theoretical rescaling procedure is that the relationship between A and $\tilde{A}$ is linear. The slope of this linear relationship is symbolized as b_3, and the intercept as a_3.

$$A = b_3\tilde{A} + a_3. \tag{7}$$

Substituting for A in Equation 2 using Equation 7, and then solving for $\tilde{A}$ gives

$$\tilde{A} = \frac{A_r R_r}{b_3}\cdot\frac{1}{R} - \frac{a_3}{b_3}. \tag{8}$$

If the slope, b_1, and intercept, a_1, of Equation 8 are defined as

$$b_1 = \frac{A_r R_r}{b_3}, \text{ and} \tag{9}$$

$$a_1 = -\frac{a_3}{b_3}, \text{ then} \tag{10}$$

$$\tilde{A} = b_1\cdot\frac{1}{R} + a_1. \tag{11}$$

Therefore, the relationship between the measure of percent comprehension, $\tilde{A}$, and 1/R is linear with a slope of b_1 and an intercept of a_1 when $t \leq t_r$. When $t \geq t_r$, then substituting for A in Equation 4 using Equation 7 and solving for $1/\tilde{A}$ R

gives

$$\frac{1}{\tilde{A}_R} = \frac{b_3}{1-a_3-a_3 i R}\frac{1}{R} + \frac{b_3 i}{1-a_3-a_3 i R} \tag{12}$$

Defining $\tilde{A}_R$ as the observed efficiency score, $\tilde{E}$, and substituting for $\tilde{A}_R$ in Equation 12 gives

$$\frac{1}{\tilde{E}} = \frac{1}{1-a_3-a_3 i R}\frac{1}{R} + \frac{b_3 i}{1-a_3-a_3 i R} \tag{13}$$

Thus, Equation 13 expresses a relationship between the efficiency score, $\tilde{E}$, and rate, R, in a manner somewhat similar to Equation 3. However, it may be noted that Equation 13 is not a linear relationship. The multiplier of 1/R is not a constant, as it is in Equation 3, so it cannot be the slope of the straight line and the final term is not constant, as it is in Equation 3, so it cannot be the y-intercept. The slope and intercept would be constants were it not for the value of $a_3 iR$ which of course varies with R.

Although the relationship between $1/\tilde{E}$ and 1/R is not perfectly linear, as indicated by Equation 13, it turns out that for the range of values that are found in actual reading situations, the relationship is in fact almost perfectly linear. Therefore, the assumption of linearity is an extremely good assumption; the degree to which this is a valid assumption will be discussed later. The relationship between $1/\tilde{E}$ and 1/R is assumed to be as follows:

$$\frac{1}{\tilde{E}} = b_2\frac{1}{R} + a_2, \text{ or} \tag{14}$$

$$\frac{1}{\tilde{A}_R} = b_2\frac{1}{R} + a_2 \text{ when } t \geq t_r. \tag{15}$$

Two points define a straight line, with a corresponding slope and intercept. Therefore, when there are at least two values of $\tilde{A}$ at two reading times less

than t_r, and at least two more values of $\widetilde{A}$ at two reading times greater than t_r, then the slopes and intercepts of the curve constants in Equations 11 and 14 are determinable. This means that b_1, a_1, b_2, and a_2 can be calculated when there are at least 4 values of $\widetilde{A}$.

Equation 15 can be solved for $\widetilde{A}$ to give an equation that is similar to Equation 5,

$$\widetilde{A} = \frac{1}{b_2 + a_2 R} \tag{16}$$

When $t = t_r$, then $R = R_r$ and A_o is the same value for both Equations 11 and 16. Substituting R_r for R in both Equations 11 and 16 and setting their right terms equal to each other, since both are equal to $\widetilde{A}$ when $t = t_r$, yields the following equation:

$$\frac{1}{b_2 + a_2 R_r} = b_1 \frac{1}{R_r} + a_1 \tag{17}$$

This equation can be rearranged to equal zero, as follows:

$$a_1 a_2 R_r^2 + (a_2 b_1 + a_1 b_2 - 1) R_r + b_1 b_2 = 0 \tag{18}$$

Equation 18 is the familiar quadratic equation which has the following solution:

$$R_r = \frac{-(a_2 b_1 + a_1 b_2 - 1) - \sqrt{(a_2 b_1 + a_1 b_2 - 1)^2 - 4(a_1 a_2)(b_1 b_2)}}{2 a_1 a_2} \tag{19}$$

The next value to be solved for is A_r. To achieve this, three equations will be developed which have three unknowns, one of which is A_r. Then, the three equations will be solved simultaneously to provide a solution for A_r.

Substituting for i in Equation 4, using Equation 6, and letting $R = R_r/2$, yields

$$\frac{1}{A_{\frac{R_r}{2}}} = \frac{1}{R_{\frac{r}{2}}} + \left(\frac{1}{A_r} - 1\right). \tag{20}$$

Equation 20, above, can be solved for A to give

$$A = \frac{2A_r}{A_r + 1}, \tag{21}$$

the first of the three equations.

For the next equation, the first step will be to solve Equation 7 for A_o,

$$\tilde{A} = \frac{1}{b_3}A - \frac{a_3}{b_3}, \tag{22}$$

and then substitute into Equation 22 using Equation 10 to give

$$\tilde{A} = \frac{1}{b_3}A + a_1. \tag{23}$$

Substituting for $\tilde{A}$ in Equation 16, using Equation 23, and letting $R = R_r/2$ gives

$$\frac{1}{b_3}A + a_1 = \frac{1}{b_2 + a_2 R_{\frac{r}{2}}}, \tag{24}$$

the second equation.

The third equation is Equation 9,

$$b_1 = \frac{A_r R_r}{b_3} \tag{25}$$

In the three designated equations above, the three unknowns are A_r, b_3, and A. A being the accuracy when the rate is only one-half as fast as the rauding rate, i.e., $R_r/2$, and the time per thought, $2/R_r$, is twice as long as the time

required to read the passage once.

Solving these three equations simultaneously gives the following solution for A_r:

$$A_r = \frac{2b_1}{\frac{2R_r}{2b_2 + a_2R_r} - a_1R_r} - 1. \qquad (26)$$

Thus, when a_1, b_1, a_2, and b_2 are known, then R_r can be solved for using Equation 19, and A_r can be solved for using Equation 26. Then, the value of b_3 can solved for using Equation 25, and in turn, a_3 can be solved for using Equation 10. There is no reason to solve for A because the value of A for $R = R_r/2$ is not of general value; it was only used for convenience in solving for A_r which in turn allows the remaining curve parameters, b_3 and a_3, to be solved for.

In summary,

(a) when it is assumed that the relationship between $1/\tilde{A}$ R and 1/R is linear,

(b) when it is also assumed that the relationship between $\tilde{A}$ and A is linear, and

(c) when two or more percent comprehension scores, $\tilde{A}$, are available at two times below t_r, and two or more other percent comprehension scores, $\tilde{A}$, are available at two times above t_r,

then the two most important curve parameters in rauding theory can be calculated, A_r and R_r, as well as all of the other less important parameters such as a_1, b_1, a_2, b_2, a_3, and b_3.

As long as A_r and R_r are known, then values of A at any rate can be estimated. Using Equations 2 and 5, this means that the values of $\tilde{A}$ at any particular rate, R, can be transformed into estimated values of A. Once the values of b_3 and a_3 are known, then any value of $\tilde{A}$ can be substituted into Equation 7 and a rescaled value of A can be solved for.

This completes the outline of the theoretical procedures for the absolute rescaling of the percent comprehension scores, $\tilde{A}$.

Validity of the Theoretical Rationale

Linearity of A versus 1/R

Equation 2 reflects the relationship between A and 1/R as prescribed in rauding theory. By substituting for A in Equation 1, using Equation 2, it can be seen that $E=A_r R_r$ when $t \leq t_r$. Thus, one of the deductions that can be made from this part of the theory is that AR or E should be constant for an individual who has not been given enough time to finish reading a passage. The data contained in the original presentation of rauding theory (Carver, 1977a) supports the contention that E is constant when the average rate of reading is likely to be less than R_r. Direct empirical support for the theoretical relationship in Equation 2 being linear, the slope being $A_r R_r$, and the y-intercept being zero, comes from a study of reading rates varying from 62.5 Wpm to 100,000 Wpm (Carver, 1978b). Thus, it appears that there is empirical support for Equation 2, one of the two most important theoretical relationships in rauding theory.

Linearity of 1/E versus 1/R

Equation 3 reflects the relationship between 1/E and 1/R when $t \leq t_r$ as prescribed in rauding theory. This part of the theory was developed from empirical data (Carver, 1977-78). Subsequent to that, extensive data involving college students reading material at four levels of material difficulty and at rates from 83 Wpm to 500 Wpm provided strong support for the linearity of this relationship (Carver, 1978a). Thus, it appears that there is also empirical support for Equation 3, the second of the two most important theoretical relationships in rauding theory.

Linearity of A versus $\tilde{A}$

Equation 7, or Equation 22, reflects an assumption or condition that expands reading theory, i.e., that the relationship between the theoretical construct A and the empirical indicator of the accuracy of comprehension, $\tilde{A}$ is linear.

Of course, not all conceivable percent comprehension measures, or empirical indicants, would satisfy this assumption. An empirical measure of how much of a passage an individual has comprehended would start at a certain point near zero, 0% comprehension, when the individual began to read the passage and would increment linearly with the amount of time spent reading until finishing where comprehension would be near 1.00, or 100% comprehension. Of course, it is easy to conceive of measures which would not have this type of a function. Some comprehension tests, for example,may focus upon the last part of a passage so that no gain in $\tilde{A}$ occurs as A is increasing initially. Conversely, some measures may focus upon the information that is acquired initially so that $\tilde{A}$ jumps precipitously at the outset and does not increase thereafter while A is increasing steadily from the beginning to the end.

It would seem to be relatively easy to determine the likelihood of certain comprehension measures meeting this assumption. Those measures which are based upon algorithms which sample the text of a passage at regular intervals would seem to satisfy this assumption. The chunked test and the cloze test used in the research presented by Carver (1973) would seem to satisfy this criterion, and indeed, the relationship reported between these objective types of tests and a subjective measure of understanding was found to be linear in that research. The reading-storage algorithm for making comprehension tests (Carver, 1975a) would also seem to satisfy this criterion because the number of test items are linearly related to the amount of text. The data presented by Carver (1978a) using this reading-storage test provides strong empirical support for reading theory and it is unlikely that this could happen without there being a linear relationship between

this empirical indicator, $\tilde{A}$, and A. Another type of measure that has been used in past research involves the development of a multiple choice questions on each sentence in a passage (Carver, 1975b; Carver 1978a). Thus, one could expect $\tilde{A}$ to increase approximately linearly with A using this type of measure, and again, the support that this measure reflected for rauding theory (Carver, 1978a) would indirectly provide support for its linear relationship with A.

In general, when the measure of comprehension employed has been developed in a manner such that the number of test items is approximately linearly related the number of words in the passage, then it would seem that the assumption is reasonably valid that the relationship between the percent comprehension scores, $\tilde{A}$, is linearly related to the accuracy of comprehension A. For other techniques that are used to measure comprehension, the validity of this particular assumption may be highly questionable.

Linearity of $\tilde{A}$ versus 1/R

As shown in the development of Equation 8, the assumption of linearity between $\tilde{A}$ and 1/R must be valid, given the validity of the assumed linear relationship between A and 1/R and given the validity of the linear relationship between A and $\tilde{A}$.

Linearity of $1/\tilde{E}$ versus 1/R

As noted earlier from Equation 13, the relationship between $1/\tilde{E}$ and 1/R is definitely not linear given the validity of the proceeding assumptions. However, it turns out that the value of a_3/R in Equation 13 is relatively small and inconsequential in most of the reading situations to which it is desirable to generalize so that for all practical purposes this relationship is perfectly linear. This can be demonstrated. Table 1 contains values for six cases that represent the range of reading situations that are likely to be encountered in the real world. Cases 1, 2, and 3 represent college level readers, $R_r = 18$ (a reader that reads 18 standard sent-

ences per minute or 300 standard length words per minute), and cases 4 - 6 represent a second grade reader, $R_r = 9$ (a reader that reads 9 standard sentences per minute or 150 standard length words per minute). Cases 1 and 4 represent situations where the material being read is difficult, $A_r = .22$, i.e., the difficulty of the material is about three grade levels higher than the ability of the material (see Table 3, Carver 1977-78). Cases 2 and 5 represent situations where the difficulty of the material is approximately equal to the ability of the individual so that $A_r = .50$ (see Table 3, Carver 1977-78). Cases 3 and 6 represent situations where the difficulty of the material is about six grade levels lower than the ability of the individual, $A_r = .87$ (also taken from Table 3, Carver 1977-78).

Insert Table 1 about here

The sensitivity of the test to increments in accuracy, A, is indicated by the slope of $\tilde{A}$ regressed on A, i.e., $1/b_3$ as indicated in Equation 23. The higher the slope, the more valid or sensitive the test is to changes in the accuracy of comprehension. The $\tilde{A}$ intercept, i.e., a_1 as indicated in Equation 23, reflects how well people do on the test without getting to read the passage, i.e., when A = 0. This variable has been termed "passage dependency" by other investigators (see Tuinman, 1973-74); when a_1 is low then passage dependency is high. In Cases 3 and 6 the test used was very insensitive, $1/b_3 = .2$ and it has relatively low passage dependency, $a_1 = .50$. In Cases 2 and 5, the test used was very sensitive, $1/b_3 = .90$, and it has relatively high passage dependency, $a_1 = .01$. Cases 1 and 4 represent situations in between the preceeding extreme values, i.e., $1/b_3 = .5$ and $a_1 = .3$.

The above values were used to solve for b_3, a_3, and I (using Equation 6), and then these values were substituted into Equation 13 to produce values of $1/\tilde{E}$ for each corresponding value of 1/R. For the assumption in question to hold, the average rate of presentation of the material must not be higher than the rauding

rate, i.e., $R \leq R_r$. Therefore, the values of R in Cases 1 - 6 must be equal to or less than R_r. In cases 1 - 3, R ranged from 18 down to 6 and in Cases 4 - 6, R ranged from 9 down to 5.

The six sets of values for $1/\tilde{E}$ are contained in Table 1. Since the correlation coefficient is a measure of the degree to which a straight line fits the data, this coefficient was calculated for the six relationships between $1/\tilde{E}$ and 1/R in Table 1. These results are also presented at the bottom of Table 1. Notice that the six correlations ranged from a high of 1.000 to a low of .999, rounded to three decimal places. These data suggest that the assumption of linearity between $1/\tilde{E}$ and 1/R is probably extremely valid for the range of situations normally encountered in reading where tests of comprehension are employed.

Summary

It appears that there is ample empirical and rational support for the functional relationships advanced for rescaling percent comprehension scores along an absolute scale reflecting the accuracy of comprehension.

Description of the Rescaling Procedures

This section will outline the procedures used to apply the theoretical rationale for rescaling, presented earlier.

As noted earlier, the rescaling procedure must have at least four data points to function. At least two of the data points must be observed percent comprehension scores, $\tilde{A}$, when the time allowed for reading is less than (or equal to) the time required to read the passage once at the rauding rate, i.e., $t \leq t_r$. At least two other data points must be observed percent comprehension scores, $\tilde{A}$, when the time allowed for reading is greater than (or equal to) the time required to read the passage once at the rauding rate, i.e., $t \geq t_r$. Much of the time, therefore, the first problem to be solved concerns whether $t > t_r$ or $t < t_r$. Since t_r depends upon R_r, and R_r is a result of the scaling

procedure this is a circular problem. When there are only four data points, then there is only one possible solution to this problem, albeit a poor one; the first two points must be considered as representing situations where $t \leq t_r$ and the last two points must be considered as representing situations where $t \geq t_r$.

When there are more than 4 data points then the best fit procedure is employed. This procedure fits every possible set of curves to the data and then the curve that fits the data best is selected. Before initiating the best fit procedure, the data points (consisting of paired X and Y values, letting $1/\tilde{R}$ be X and $\tilde{A}$ be Y) are ranked according to the lowest to the highest values of $1/\tilde{R}$. For example, the data point with the lowest value of $1/\tilde{R}$ would be designated as data point number 1, the next lowest would be data point number 2, and so on up to the data point with the highest value of $1/\tilde{R}$. Then, the steps for the best fit procedure are described as follows:

Step 1. Data points 1 and 2, i.e., the first two data points, which represent the two smallest $1/\tilde{R}$ values and their corresponding $\tilde{A}$ values, are considered as representing situations where $\tilde{t} \leq \tilde{t}_r$, and data points from 2 to N (let N represent the last data point) are considered as representing situations where $\tilde{t} \geq \tilde{t}_r$.

Step 2. For data points 1 and 2, the part of the theoretical scaling procedure relevant to situations where $\tilde{t} \leq \tilde{t}_r$ is appropriate. For example, the slope constant $\tilde{b}_1$ and y-intercept $\tilde{a}_1$ for Equation 11 can be estimated by simply fitting a least squares straight line to these two data points representing paired values of $\tilde{A}$ regressed on $1/\tilde{R}$.

Step 3. For data points 2 through N, the paired values of $1/\tilde{E}$ regressed on $1/\tilde{R}$ are calculated and a least squares linear fit is also made to these data. The slope and intercept of this line are therefore the

the values used for b_2 and a_2, respectively in Equation 14.

Step 4. From the above values of $\tilde{a}_1$, $\tilde{b}_1$, $\tilde{a}_2$, and $\tilde{b}_2$, the remaining uncalculated. The estimated rauding rate, $\tilde{R}_r$, is calculated using the following equation:

$$\tilde{R}_r = \frac{-(\tilde{a}_2\tilde{b}_1+\tilde{a}_1\tilde{b}_2-1) - \sqrt{\tilde{a}_2\tilde{b}_1+\tilde{a}_1\tilde{b}_2-1)^2 - 4(\tilde{a}_1\tilde{a}_2)(\tilde{b}_1\tilde{b}_2)}}{2\tilde{a}_1\tilde{a}_2} \quad (27)$$

The estimated rauding accuracy, $\tilde{A}_r$, is calculated as follows:

$$\tilde{A}_r = \frac{2\tilde{b}_1}{\frac{2\tilde{R}_r}{2\tilde{b}_2 + \tilde{a}_2\tilde{R}_r} - \tilde{a}_1\tilde{R}_r} - 1 \quad (28)$$

A modification of Education 25 is used to calculated $\tilde{b}_3$.

$$\tilde{b}_3 = \frac{\tilde{A}_r\tilde{R}_r}{\tilde{b}_1} \quad (29)$$

A modification of Equation 10 is used to calculated $\tilde{a}_3$.

$$\tilde{a}_3 = -\tilde{a}_1\tilde{b}_3. \quad (30)$$

Step 5. Derived from Equation 11, the following equation is used to calculate values of $\tilde{A}'$ for each value of $\tilde{R}$ at data points 1 and 2.[3]

$$\tilde{A}' = \tilde{b}_1 \cdot \frac{1}{\tilde{R}} + \tilde{a}_1. \quad (31)$$

Step 6. Derived from Equation 16, the following equation is used to calculate values of $\tilde{A}'$ for each value of $\tilde{R}$ at data points 2

through N.

$$\tilde{A}' = \frac{1}{\tilde{b}_2 + \tilde{a}_2 \tilde{R}} . \tag{32}$$

Step 7. The Average Empirical Error (AEE) of estimate is developed by calculating $\tilde{A} - \tilde{A}'$ for each $1/\tilde{R}$ value and then calculating the root mean square (RMS) of the errors for all the data points. There will be N+1 errors because point 2, the pivitol point, gets used twice for AEE.

Step 8. The $\tilde{A}$ values at each $\tilde{R}$ can be rescaled into values of A, designated as $\bar{A}$, using the values of $\tilde{b}_3$ and $\tilde{a}_3$ that resulted from the best fit,[4] and substituting them into the following equation which is derived from Equation 7:

$$\bar{A} = \tilde{b}_3 \tilde{A} + \tilde{a}_3 . \tag{33}$$

The Average Theoretical Error (ATE) of estimate is developed by calculating $\bar{A} - \bar{A}'$ for each value of $1/\tilde{R}$ and then calculating the RMS of all N of these errors. The $\bar{A}'$ values for each value of $1/\tilde{R}$ can be calculated by substituting the $\tilde{A}_r$ and $\tilde{R}_r$ values into Equations 2, 5, and 6, i.e.,

$$\bar{A}' = (\tilde{A}_r \tilde{R}_r) \frac{1}{\tilde{R}} \text{ when } t \leq t_r \text{ or } \tilde{R} \geq \tilde{R}_r ; \tag{34}$$

$$\bar{A}' = \frac{1}{1 + \frac{1}{\tilde{R}_r} \left(\frac{1}{\tilde{A}_r} - 1 \right) \tilde{R}} \text{ when } t \geq t_r \text{ or } \tilde{R} \leq \tilde{R}_r . \tag{35}$$

Step 9. The Average Error (AE) of estimate is developed by calculating the RMS of all of the N+1 errors for AEE and the N errors for ATE, i.e., the RMS of the 2N+1 errors calculated in Step 7 and Step 8.

Step 10. Steps 1-9 are repeated, but this time data points 1-3 represent the $t \leq t_r$ situations and data points 3 through N represent the $t \geq t_r$ situations. For example, for data points 1-3, at least squares linear fit is made to the $\tilde{A}$ regressed on $1/\tilde{K}$ data, and the resulting slope and intercepts are the values of $\tilde{b}_1$ and $\tilde{a}_1$. After the appropriate procedures described above are applied to data points 3-N, the AEE, ATE, and AE values are again calculated.

Step 11. Step 10 continues to be repeated by incrementing the pivitol dividing point, e.g., point 3 in Step 10, one unit each time until this pivitol dividing point equals N-1. For this last fit, there are only two data points for the $t \geq t_r$ fit so the least squares procedure is not used to calculate $\tilde{a}_2$ and $\tilde{b}_2$ just as it was not used in Step 2 to calculate $\tilde{a}_1$ and $\tilde{b}_1$.

Step 12. The best fit of all the N-2 fits is the one with the smallest AE value, and the curve parameters from this particular fit, e.g., $\tilde{a}_1$, $\tilde{b}_1$, $\tilde{a}_2$, $\tilde{b}_2$, $\tilde{a}_3$, $\tilde{b}_3$, $\tilde{R}_r$, and $\tilde{A}_r$, values, are considered to be the best estimates available

The above 12 steps provide an objective technique for rescaling the scores on an empirical measure of comprehension so that they are on an absolute scale from 0 to 100 representing percentage units of the accuracy of comprehension. These rescaled values, $\tilde{A}$, may also be used to calculate an estimate of the absolute number of thoughts in the passage that were comprehended by using equations given in the original presentation of the theory (Carver, 1977-78).

The degree to which the best fitting solution fits the data can be measured on an absolute scale by the ATE value.

The presentation of the rescaling steps in this section would have been complete at this point were it not for subsequent applications of these steps to actual and hypothetical data. This pilot testing stimulated some modifications of the procedures.

Three types of undesirable events occurred in the pilot testing. The first involved imaginary solutions, i.e., the square root of a negative number in the quadratic equation--Equation 27 solving for R_r. This undesirable soltuion occurs when the curve fit to the data points where $t \leq t_r$ does not cross the curve fit to the data points where $t \geq t_r$; theoretically, they must cross at $\tilde{R}_r$. The second undesirable solution involved values of $\tilde{a}_2$ which were less than zero. This has the effect of providing values of $\tilde{A}$ which decrease as $1/\tilde{R}$ increases, a theoretically unacceptable soltuion. When $\tilde{b}_2$ is less than zero, it is possible for $\tilde{A}$ to be negative, another theoretically unacceptable solution. The modification procedures presented below seem to have eliminated these undesirable types of solutions:

(a) when the value of the term under the radical in Equation 27 is less than zero, the value of this term is set equal to zero.

(b) when either $\tilde{a}_2$ or $\tilde{b}_2$ is initially calculated in Steps 2, 10, or 11, and is less than zero, i.e., $\tilde{a}_2 < 0$, or $\tilde{b}_2 < 0$, then $\tilde{b}_2$ is set equal to 1.00 and $\tilde{a}_2$ is derived by fitting a line with a slope of 1.00 through the $1/\tilde{E}$ and $1/\tilde{R}$ data points. An intercept value is calculated for a line fit through each point with a slope at 1.00 and the new value of $\tilde{a}_2$ is the mean of all these intercept values.

Applicetion of the Rescaling Procedures

In a study of reading comprehension at normal and speed reading rates (Carver, 1978b), passages were presented for varying length of times with corresponding

average rates, $\tilde{R}$, of presentation. Four different measures of comprehension were administered subsequent to the presentation of the passages. The four measures may be considered as providing different observed indicants of the accuracy of comprehension, $\tilde{A}$, at each of the administered rates, $\tilde{R}$. Each measure tended to provide different percent comprehension values, $\tilde{A}$, at each rate, $\tilde{R}$, even though theoretically there should only be one value of accuracy of comprehension, A, at each rate. Thus, these data seemed appropriate for evaluating whether the rescaling procedure described earlier in this article tends to provide approximately the same four values of $\bar{A}$, or $\bar{A}'$, at each rate even though the four values of $\tilde{A}$ were different at each rate.

A complete description of the original experimental methods is contained elsewhere (Carver, 1978b). Only those aspects of the earlier research which are directly relevant to the present analysis will be described in detail.

There were 10 rates employed in the original research, 62.5, 125, 250, 500, 1,000, 2,000, 4,000, 8,000, 16,000, and 100,000 wpm. There were 10 different, 100 word passages at the tenth grade level of difficulty and the order of presentation of the passages and the rates was counter-balanced. Rate was not directly controlled, i.e., the amount time the passage was presented on the screen using a motion picture projector, was directly controlled. There were a total of 10 different films, each containing 10 passages at 10 different rates. A total of 102 subjects were tested in 20 groups.

The four different measures were Understanding Judgments, Missing Verbs, Best Titles Test, and Sentence-Halves Test. The Understanding Judgments measure was simply the subjects opinion regarding what percent of the passage he or she had understood. At each rate the mean of these judgments was calculated. For the Missing Verbs measure, a verb was deleted in certain sentences. In one-half of the passages, five, a verb had been deleted from one of the sentences in

the passage; one passage had the verb missing in the first 1/5 of the passage, another passage had the verb deleted in the second 1/5 of the passage, etc., so that the location of the missing verbs were equally spaced through the passages. Immediately after a passage had been presented, the subject circled yes or no indicated whether the subject thought a verb was missing in the passage. The percent of individuals with correct answers, corrected for guessing, at each rate was the Missing Verbs measure. The Best Titles Test, administered after all 10 passages had been presented, consisted of 10 multiple-choice questions. Each question consisted of five possible titles for a passage and the subjects task was to select the best title. The percent of questions answered correctly, after a correction for guessing had been applied, constituted the Best Titles measure. The Sentence-Halves Test contained 40 sentence-halves, 20 were halves of sentences taken from the 10 passages and 20 were halves of sentences taken from passages that had not been presented. The subjects were to answer yes or no as to whether they thought each sentence-half had or had not been contained in the 10 passages they had just been presented. Again, after a correction for guessing had been applied, this measure was the mean of the percent correct at each rate.

Fig. 1 contains the resulting data; $\tilde{A}$ for each of the four measures has been plotted as a function of $1/\tilde{R}$, min. per sentnece. The highest $1/\tilde{R}$ values represent the lowest rates, i.e., 62.5 and 125 Wpm. The values of $\tilde{A}$ for the five highest rates, 2,000 to 100,000 Wpm were omitted, for two reasons: (a) all 20 of the values lie in the narrow range of $1/\tilde{R}$ values between .00 and .01 and this tends to clutter the graph, and (b) the $\tilde{A}$ values fluxuate inconsistently around zero for the three objective measures because chance is likely the primary factor operating at these rates.

Insert Fig. 1 about here

The data in Fig. 1 vividly portray how different measures of comprehension yield different absolute percent scores. The Understanding Judgments measure and the Missing Verbs measure tend to start near zero when the $1/\tilde{R}$ value is near zero and both increase to near perfect $\tilde{A}$ values when the $1/\tilde{R}$ value is high. Both the Best Titles Test and the Sentence-Halves Test appear to be difficult tests since the $\tilde{A}$ values do not get above about .50 even at the high $1/\tilde{R}$ values. Also included in Fig. 1 are 5 hypothetical data points which will be explained later.

The data in Fig. 1 was rescaled using the rauding rescaling procedure described earlier in Steps 1-12. These rescaled data are presented in Fig. 2, i.e., $\bar{A}$ is plotted as a function of $1/\tilde{R}$ for each of the four measures. It may be noted that all four measures tend to produce approximately the same $\bar{A}$ values. The rauding rescaling formulas, derived from Equation 33, are as follows (with the values of $\tilde{b}_3$ and $\tilde{a}_3$ rounded to two decimals):

(a) for Understanding Judgments,

$$\bar{A} = .94\,\tilde{A} + .00,$$

(b) for Missing Verbs

$$\bar{A} = .86\,\tilde{A} + .14,$$

(c) for Best Titles Test,

$$\bar{A} = 2.45\,\tilde{A} - .39,$$

and (d) for Sentences-Halves Test,

$$\bar{A} = 1.64\,\tilde{A} + .14.$$

The $\tilde{A}$ values from the three objective measures, Missing Verbs, Best Titles Test, and Sentence-Halves Test, are examples of measures that are commonly re-

ferred to as "percent comprehension scores," i.e., when the $\widetilde{A}$ values are converted from proportions to percents by multiplying by 100. Therefore, it can be seen that the percent comprehension scores in Fig. 1 yield quite different results while their rescaled counterparts in Fig. 2 yielded almost exactly the same results for rescaled accuracy of comprehension, $\bar{A}$. In order to quantify the results visable in Fig. 1 and Fig. 2, average errors or differences were calculated. For example, at the highest $1/\widetilde{R}$ value, $1/\widetilde{R}$ = .267, there are four A values, .99, .92, .53, and .48. The six absolute differences between each pair of values are .07, .46, .51, .39, .44, .05. These differences were calculated for all 5, $1/\widetilde{R}$ rates, and the mean of these 30 differences, or errors, was .22. The corresponding mean for the $\bar{A}$ data in Fig. 2 is only .05. The difference between .22 and .05 partially summarizes the effect of the rescaling procedure.

Insert Fig. 2 about here

For someone with only a superficial knowledge of the rescaling procedure, it may be suspected that somehow the rescaling procedure artifactually makes the curves go up to near 1.00 values. To illustrate how the rescaled values do not necessarily have to approach 1.00, the hypothetical data presented in Fig. 1 were also rescaled and the results are presented in Fig. 2. It may be noted that these values do not coincide with the other curves.

All the curve parameters that resulted from the application of the rauding rescaling procedure are presented in Table 2. The hypothetical data has been presented for completeness but will be disregarded in the discussion. The best fitting curves for each measure was found when the pivitol data point was either data point No. 2 or No. 3. The estimated rauding rate for the four measures varied from 238 to 481 Wpm and the estimated rauding accuracies varied from .62 to .77.

The test sensitivities varied from a low of .41 for the Best Titles Test to a high of 1.17 for the Missing Verbs test. The average errors were all quite small; the Average Theoretical Errors varied from .01 to .03 suggesting all four of the best fits were excellent fits.

Insert Table 2 about here

One of the most amazing aspects of the data is the high correspondence between the values for the Understanding Judgments and the Missing Verbs measures. The average of their differences, $\bar{A} - \bar{A}'$, at each rate was only .01. These two measures are derived in quite different ways. The Understanding Judgments measure consists of a mean of a group of judgments, e.g., a value of .30 would indicate that judgments at this rate would probably range from about 10 percent to 50 percent but the mean of all the percents was 30. If the Missing Verbs measure had a corresponding 30 value at the same rate, this would mean that 65 percent of the individuals had correctly answered whether there had in fact been a verb missing in the passage they had just been presented--65 percent corrected for guessing is 65-35 or 30 percent. A rescaling procedure that gives approximately the same rescaled values from an opinion poll, such as Understanding Judgments, and an objective measure, such as Missing Verbs, suggests that it is a relatively accurate and valid procedure.

Another way to evaluate the procedure is to make the objective measures as reliable, and therefore as valid, as possible. This can be accomplished by getting the mean of the $\bar{A}$ values derived from the three objective measures at each rate. This has been done and the results are presented in Fig. 3. Also in Fig. 3 are the data points for the original Understanding Judgments, i.e., the $\widetilde{A}$ values. The data points at each rate are very close together. The mean of the five ab-

solute differences at each rate is only .05, a relatively small difference. The continuous curve presented in Fig. 3 resulted from another estimation procedure. From the median score the subjects made on a standardized reading test that original presentation of rauding theory (Carver, 1977-78). These values were then substituted into Equations 34 and 35, and the results have been presented in the form of the continuous curve in Fig. 3. Note that these are $\hat{A}$ values which are not rescaled values at all.[5] The coincidence of these three methods of estimating absolute amounts of the accuracy of comprehension is phenominal. The A values on the continuous curve differ on the average only .04 from Understanding Judgements and only .02 from the combined objective measures. Again, these data suggest validity for both the rauding rescaling procedure as well as the theory which was used to generate the procedure.

It should be pointed out that the results presented in connection with the original presentation of these same data (Carver, 19;6b) provided similar results. However, there is an extremely important difference. In that earlier analysis, a linear relationship was assumed between the Understanding Judgements and each of the other three objective measures; a least squares fit with resulting slopes and intercepts was used to rescale the objective measures so that they had the same mean and standard deviation as the Understanding Judgements. In order for that type of analysis to be repeated, it would be necessary to always have Understanding Judgments and the validity of the analysis would always be directly dependent upon the validity of the Understanding Judgments. It should not go unnoticed that the $\bar{A}$ values for each of the three objective measures in the present analysis were derived independent of the Understanding Judgments variable. This means that research can be conducted which uses only one measure which can reasonably be assumed to meet the assumption of being linearity related to A. With only one measure and one reasonably valid assumption, the measure can be rescaled

to reflect absolute amounts of the accuracy of comprehension, $\bar{A}$.

Summary and Conclusions

Use of reading comprehension tests often result in what are popularly called "percent comprehension" measures. The absolute sizes of the percent comprehension values depend upon who made the test, and how it was made. The absolute percent comprehension found in the research cannot reasonably be generalized to another type of test or another researcher who tries to make the same type of test. For example, 90 percent comprehension score, $\tilde{A} = .90$, may reflect almost total comprehension, e.g., $A = .90$, or it may simply indicate that extremely easy test questions were used while comprehension of the passage was actually low, e.g., $A = .10$. Ordinarily it is impossible to determine what A really was, given $\tilde{A}$. The rauding rescaling procedure described in this article provides a remedy for this. This procedure allows the percent comprehension measure, symbolized by $\tilde{A}$, to be rescaled so that it reflects absolutes amounts of a theoretical variable called accuracy of comprehension and symbolized by $\bar{A}$. This rescaling of percent comprehension involves a series of theoretical formulas derived from a theory of reading comprehension, called rauding theory. In order for the rescaling procedure to work, the percent comprehension measure must meet one primary assumption. This assumption is that the relationship between $\tilde{A}$ and A is linear. It is reasonable to make this assumption any time that the number of items on the test is linearly related to number of words in the reading passage.

The rauding rescaling procedure was applied to research data that were originally presented elsewhere (Carver, 1978b). In that previous research, four different measures of percent comprehension were administered after an individual had read a passage at a certain rate. The four measures provided quite different values of $\tilde{A}$, even though it was logical to expect that all four should be reflecting exactly the same amount of the accuracy of comprehension variable, A. When the rauding rescaling procedure was applied to the four measures, all four provided approximately the same values of $\bar{A}$, the rescaled values of the accuracy of comprehension. Thus, there is evidence that this procedure provides a valid technique for use with measures of percent comprehension derived from objective tests so that their rescaled measures reflect absolute amounts of the accuracy of comprehension.

REFERENCES

Carver, Ronald P. Understanding, information processing and learning from prose materials. Journal of Educational Psychology, 1973, 64, 76-84.

Carver, Ronald P. Measuring the primary effect of reading: Reading storage technique, understanding, judgments, and cloze. Journal of Reading Behavior, 1974, 6, 249-274.

Carver, Ronald P. Revised procedures for developing reading-input materials and reading-storage tests. Journal of Reading Behavior, 1975, 7, 155-172, (a)

Carver, Ronald P. Comparing the reading-storage test to the paraphrase test as measures of the primary effect of reading. Journal of Educational Psychology, 1975, 67, 274-284. (b)

Carver, Ronald P. Toward a theory of reading comprehension and rauding. Reading Research Quarterly, 1977-78, 13 8-63. (a)

Carver, Ronald P. Optimal rate of reading. Unpublished manuscript, 1978. (a)

Carver, Ronald P. Reading comprehension under normal and speed reading conditions, Unpublished manuscript, 1978. (b)

Tulman, J. Jaap. Determining the passage dependency of comprehension questions in five major tests. Reading Research Quarterly, 1973-74, 9, 206-223.

Footnotes

1. The preparation of the manuscript was supported in part by the National Institute of Education, Grant No. NEG-00-3-0043, and in part by the U.S. Office of Naval Research, Personnel and Training Research Programs.
2. A tilda ($\sim$) placed over the symbol for a variable will indicate that an empirical measure or procedure has been used to derive estimated values for that variable.
3. A prime ($'$) placed to the upper right side of the symbol for a variable will indicate that linear functions have been fit to the data points and the values of the variable have been derived from this prediction procedure.
4. A bar ($-$) placed over the symbol for a variable will indicate that the values for the variable have resulted from the rescaling procedure.
5. A hat ($\wedge$) placed over the symbol for a variable will indicate that the values for the variable resulted from standardized estimates of the reading ability of the individual and the reading difficulty of the material, i.e., $\widetilde{R}_r$ and $\widetilde{A}_r$ were determined independently of four or more pairs of $1/\widetilde{R}$ and $\widetilde{A}$ values.

TABLE 1

SIX HYPOTHETICAL CASES FOR INVESTIGATING THE VALIDITY OF ASSUMING A LINEAR RELATIONSHIP BETWEEN

$1/\tilde{E}$ AND $1/R$

Given	Case 1	Case 2	Case 3	Given	Case 4	Case 5	Case 6
R_r	18	18	18	R_r	9	9	9
A_r	.22	.50	.87	A_r	.22	.50	.87
$1/b_3$	.50	.90	.20	$1/b_3$	.50	.90	.20
a_1	.30	.01	.50	a_1	.30	.01	.50
Solved				Solved			
b_3	2.000	1.111	5.000	b_3	2.000	1.111	5.000
a_3	-.600	-.011	-2.500	a_3	-.600	-.011	-2.500
i	.197	.056	.008	i	.394	.111	.017
Rate R	Rate 1/R	$1/\tilde{E}$	$1/\tilde{E}$	$1/\tilde{E}$	Rate R	Rate 1/R	$1/\tilde{E}$	$1/\tilde{E}$	$1/\tilde{E}$
18	.056	.136	.121	.082	9	.111	.271	.241	.165
15	.067	.156	.134	.098	8	.125	.297	.257	.185
12	.083	.186	.152	.122	7	.143	.330	.277	.210
9	.111	.231	.183	.162	6	.167	.371	.303	.245
6	.167	.315	.244	.241	5	.200	.427	.340	.292
Correlation Coefficient		.999	1.000	1.000	Correlation Coefficient		1.000	1.000	1.000

TABLE 2

PARAMETERS FOR THE BEST FITTING CURVES

Parameter	Understanding Judgments	Missing Verbs	Sentence Halves	Best Titles	Hypothetical Data
Best Fit Point	3	3	2	2	3
a_1	.00	-.16	-.08	.16	.30
b_1	12.36	12.85	9.88	7.29	3.60
a_2	.02	.03	.06	.02	.06
b_2	.94	.98	1.84	1.82	1.03
Rauding Rate, $\widetilde{R}_r$	254	238	416	481	234
Rauding Accuracy, $\widetilde{A}_r$	.76	.77	.65	.62	.29
a_3	.00	.14	.14	-.39	-.34
b_3	.94	.86	1.64	2.45	1.14
Test Sensitivity, $1/b_3$	1.07	1.17	.61	.41	.88
AEE $(\widetilde{A} - \widetilde{A}')$	.01	.02	.02	.00	.01
ATE $(\overline{A} - \overline{A})$	.01	.02	.03	.01	.02
Average Error	.01	.02	.03	.01	.01

Figures

Fig. 1. Percent comprehension score, $\tilde{A}$, as a function of the average amount of time allowed to read each standard length sentence, $1/\tilde{R}$, in minutes per sentence, for each of the indicants of comprehension.

Fig. 2. Rescaled accuracy of comprehension, $\bar{A}$, as a function of the average amount of time allowed to read each standard length sentence, $1/\tilde{R}$, in minutes per sentence, for each of the indicants of comprehension.

Fig. 3. Accuracy of comprehension as a function of the average amount of time allowed to read each standard length sentence, $1/\tilde{R}$, in minutes per sentence, for the Understanding Judgments, $\tilde{A}$, the combination of the three rescaled objective test measures, $\bar{A}$, and the theoretically predicted values based upon reading ability and material difficulty, $\hat{A}$.

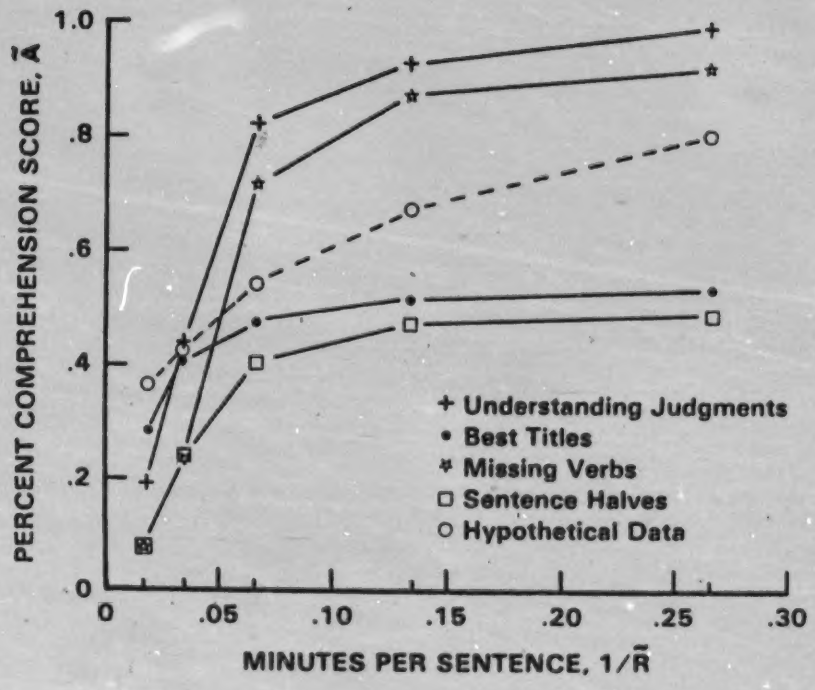

PERCENT COMPREHENSION SCORE, Ã
1.0
.8
.6
.4
.2
0
0
.05
.10
.15
.20
.25
.30
MINUTES PER SENTENCE, 1/R̃
+ Understanding Judgments
• Best Titles
* Missing Verbs
□ Sentence Halves
○ Hypothetical Data

FIG. 1

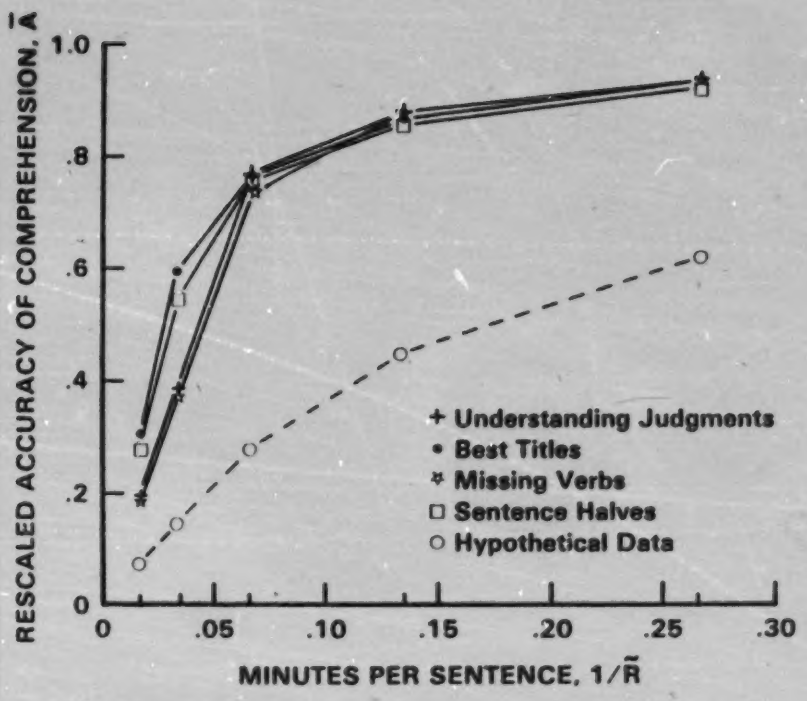

RESCALED ACCURACY OF COMPREHENSION, Ā
1.0
.8
.6
.4
.2
0
0
.05
.10
.15
.20
.25
.30
MINUTES PER SENTENCE, 1/R̃
+ Understanding Judgments
• Best Titles
✲ Missing Verbs
□ Sentence Halves
○ Hypothetical Data

FIG. 2

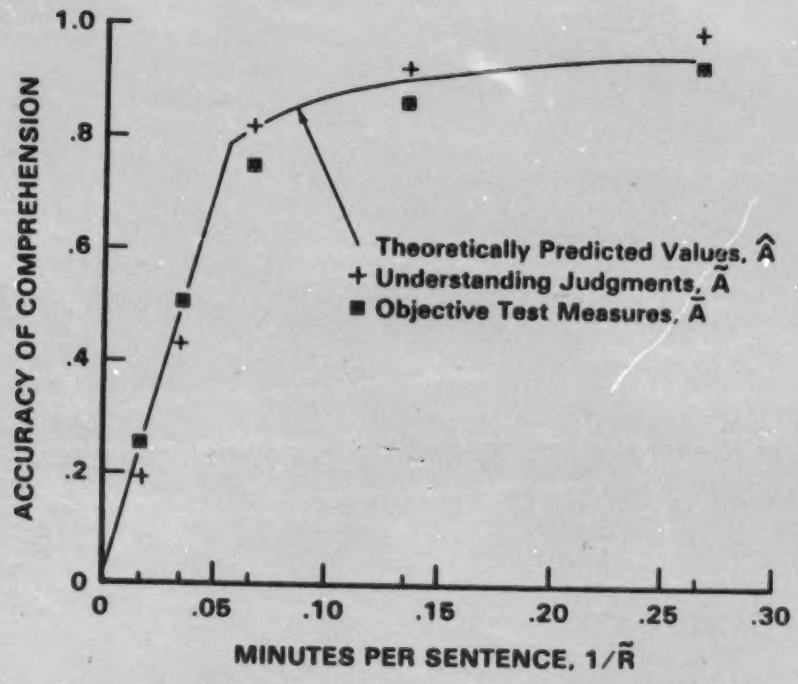

1.0
.8
.6
.4
.2
0
ACCURACY OF COMPREHENSION
0
.05
.10
.15
.20
.25
.30
MINUTES PER SENTENCE, 1/R̃
Theoretically Predicted Values, Â
+ Understanding Judgments, Ã
■ Objective Test Measures, Ā

FIG. 3